

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2002, 11:40:16 ; Search time 62 Seconds
(without alignments)
184.832 Million cell updates/sec

Title: PCT-US02-27855-1
Perfect score: 499
Sequence: 1 KTCYEGNGHFYRGKASTDTM.....YVQVGLKPLVQECMVHDCAD 86

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_101002: *
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT: *
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT: *
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT: *
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT: *
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT: *
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT: *
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT: *
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT: *
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT: *
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT: *
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT: *
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT: *
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT: *
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT: *
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT: *
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT: *
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT: *
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT: *
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT: *
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT: *
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT: *
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT: *
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	494	99.0	88	23	AAE16542 Human urokinase-ty
2	494	99.0	96	23	AAE16550 Human uPA kringle
3	494	99.0	135	23	AAE16545 Human urokinase-ty
4	494	99.0	143	23	AAE16549 Human uPA amino te
5	494	99.0	337	22	AAE16549 Human urokinase-ty
6	494	99.0	337	23	ABP41795 Human ovarian anti
7	494	99.0	365	16	AAE16549 Human urokinase-ty
8	494	99.0	378	11	AAW13635 Human prourokinase
9	494	99.0	386	16	AAE16549 Human urokinase-ty
10	494	99.0	389	11	AAW13636 Human prourokinase

11	494	99.0	390	16	AAE16542 Human urokinase-ty	Bifunctional uroki
12	494	99.0	390	16	AAE16542 Human urokinase-ty	Bifunctional uroki
13	494	99.0	392	16	AAE16542 Human urokinase-ty	Bifunctional uroki
14	494	99.0	392	16	AAE16542 Human urokinase-ty	Bifunctional uroki
15	494	99.0	392	16	AAE16542 Human urokinase-ty	Bifunctional uroki
16	494	99.0	392	16	AAE16542 Human urokinase-ty	Bifunctional uroki
17	494	99.0	392	16	AAE16542 Human urokinase-ty	Bifunctional uroki
18	494	99.0	392	16	AAE16542 Human urokinase-ty	Bifunctional uroki
19	494	99.0	392	16	AAE16542 Human urokinase-ty	Bifunctional uroki
20	494	99.0	392	16	AAE16542 Human urokinase-ty	Bifunctional uroki
21	494	99.0	392	16	AAE16542 Human urokinase-ty	Bifunctional uroki
22	494	99.0	392	16	AAE16542 Human urokinase-ty	Bifunctional uroki
23	494	99.0	392	16	AAE16542 Human urokinase-ty	Bifunctional uroki
24	494	99.0	392	16	AAE16542 Human urokinase-ty	Bifunctional uroki
25	494	99.0	392	16	AAE16542 Human urokinase-ty	Bifunctional uroki
26	494	99.0	392	16	AAE16542 Human urokinase-ty	Bifunctional uroki
27	494	99.0	392	16	AAE16542 Human urokinase-ty	Bifunctional uroki
28	494	99.0	392	16	AAE16542 Human urokinase-ty	Bifunctional uroki
29	494	99.0	392	16	AAE16542 Human urokinase-ty	Bifunctional uroki
30	494	99.0	392	16	AAE16542 Human urokinase-ty	Bifunctional uroki
31	494	99.0	392	16	AAE16542 Human urokinase-ty	Bifunctional uroki
32	494	99.0	392	16	AAE16542 Human urokinase-ty	Bifunctional uroki
33	494	99.0	392	16	AAE16542 Human urokinase-ty	Bifunctional uroki
34	494	99.0	392	16	AAE16542 Human urokinase-ty	Bifunctional uroki
35	494	99.0	392	16	AAE16542 Human urokinase-ty	Bifunctional uroki
36	494	99.0	392	16	AAE16542 Human urokinase-ty	Bifunctional uroki
37	494	99.0	392	16	AAE16542 Human urokinase-ty	Bifunctional uroki
38	494	99.0	392	16	AAE16542 Human urokinase-ty	Bifunctional uroki
39	494	99.0	392	16	AAE16542 Human urokinase-ty	Bifunctional uroki
40	494	99.0	392	16	AAE16542 Human urokinase-ty	Bifunctional uroki
41	494	99.0	392	16	AAE16542 Human urokinase-ty	Bifunctional uroki
42	494	99.0	392	16	AAE16542 Human urokinase-ty	Bifunctional uroki
43	494	99.0	392	16	AAE16542 Human urokinase-ty	Bifunctional uroki
44	494	99.0	392	16	AAE16542 Human urokinase-ty	Bifunctional uroki
45	494	99.0	392	16	AAE16542 Human urokinase-ty	Bifunctional uroki

ALIGNMENTS

RESULT 1	AAE16542	standard; Protein; 88 AA.
ID	AAE16542;	
AC	AAE16542;	
XX		
DT	09-APR-2002	(first entry)
XX		
DE	Human urokinase-type plasminogen activator (uPA) kringle.	
XX		
KW	Human: urokinase-type plasminogen activator; uPA; therapy; hypertension; stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder; microvascular occlusion; angiotensin disorder; pulmonary fibrosis; asthma; tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing; clotting disorder; uterine contraction disorder; respiratory disease; male impotence; adult respiratory distress syndrome.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200197752-A2.	
XX		
PD	27-DEC-2001.	
XX		
PF	13-JUN-2001; 2001WO-US18976.	
XX		
PR	20-JUN-2000; 2000US-212874P.	
XX		
PA	(UPE-) UNIV PENNSYLVANIA.	
XX		
PI	Cines DB, Higazi AA;	
XX		
DR	WPI; 2002-122240/16.	
DR	N-PSDB; AAD27075.	

XX Composition for modulating muscle cell and tissue contractility for
PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
PT comprising domains from urokinase-type plasminogen activator -
XX
PS Claim 1; Fig 1A; 117pp; English.
XX
CC The invention relates to a composition comprising one or more domains of
CC urokinase-type plasminogen activator (uPA). The composition is used to
CC modulate the contractility and angiogenic activity of a mammalian muscle,
CC endothelial cell or tissue. The composition is used for treating stroke,
CC hypotension, hypertension, atherosclerosis, heart attack, microvascular
CC occlusions, thrombotic microangiopathies, surgically induced thrombotic
CC disorders, angiotensin disorders, pulmonary fibrosis, asthma, tumour cell
CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,
CC diabetic retinopathy, wound healing, clotting disorder, uterine
CC contraction disorder, male impotence, respiratory disease or condition
CC such as asthma, adult respiratory distress syndrome, primary pulmonary
CC hypertension, microvascular thrombotic occlusion, and a disorder
CC associated with chronic intrapulmonary fibrin formation. The present
CC sequence is human urokinase-type plasminogen activator (uPA) kringle.
XX
SQ Sequence 88 AA;

Query Match 99.0%; Score 494; DB 23; Length 88;
Best Local Similarity 98.8%; Pred. No. 1.8e-41;
Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATV LQQT YHAHRSNALQLGKHNCRNPDN 60
Db 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATV LQQT YHAHRSNALQLGKHNCRNPDN 60

QY 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
Db 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86

RESULT 2
ID AAE16550 standard; Protein; 96 AA.
XX
AC AAE16550;
DT 09-APR-2002 (first entry)
XX
DE Human uPA kringle and connecting peptide.
XX
KW Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;
KW stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;
KW microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;
KW tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;
KW clotting disorder; uterine contraction disorder; respiratory disease;
KW adult respiratory distress syndrome; male impotence.
XX
OS Homo sapiens.
PN WO200197752-A2.
XX
PD 27-DEC-2001.
XX
PF 13-JUN-2001; 2001WO-US18976.
XX
PR 20-JUN-2000; 2000US-212874P.
XX
PA (UYPE-) UNIV PENNSYLVANIA.
XX
PI Cines DB, Higazi AA;
XX
DR WPI; 2002-122240/16.
DR N-PSDB; AAD27083.
XX
PT Composition for modulating muscle cell and tissue contractility for
PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,

PT comprising domains from urokinase-type plasminogen activator -
XX
PS Claim 25; Fig 1I; 117pp; English.
XX
CC The invention relates to a composition comprising one or more domains of
CC urokinase-type plasminogen activator (uPA). The composition is used to
CC modulate the contractility and angiogenic activity of a mammalian muscle,
CC endothelial cell or tissue. The composition is used for treating stroke,
CC hypotension, hypertension, atherosclerosis, heart attack, microvascular
CC occlusions, thrombotic microangiopathies, surgically induced thrombotic
CC disorders, angiotensin disorders, pulmonary fibrosis, asthma, tumour cell
CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,
CC diabetic retinopathy, wound healing, clotting disorder, uterine
CC contraction disorder, male impotence, respiratory disease or condition
CC such as asthma, adult respiratory distress syndrome, primary pulmonary
CC hypertension, microvascular thrombotic occlusion, and a disorder
CC associated with chronic intrapulmonary fibrin formation. The present
CC sequence is human urokinase-type plasminogen activator (uPA) kringle
and connecting peptide.
XX
SQ Sequence 96 AA;

Query Match 99.0%; Score 494; DB 23; Length 96;
Best Local Similarity 98.8%; Pred. No. 1.9e-41;
Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATV LQQT YHAHRSNALQLGKHNCRNPDN 60
Db 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATV LQQT YHAHRSNALQLGKHNCRNPDN 60

QY 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
Db 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86

RESULT 3
ID AAE16545 standard; Protein; 135 AA.
XX
AC AAE16545;
DT 09-APR-2002 (first entry)
XX
DE Human urokinase-type plasminogen activator amino terminal fragment (ATF).
XX
KW Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;
KW stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;
KW microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;
KW tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;
KW clotting disorder; uterine contraction disorder; respiratory disease;
KW adult respiratory distress syndrome; amino terminal fragment; ATF;
KW male impotence.
XX
OS Homo sapiens.
PN WO200197752-A2.
XX
PD 27-DEC-2001.
XX
PF 13-JUN-2001; 2001WO-US18976.
XX
PR 20-JUN-2000; 2000US-212874P.
XX
PA (UYPE-) UNIV PENNSYLVANIA.
XX
PI Cines DB, Higazi AA;
XX
DR WPI; 2002-122240/16.
DR N-PSDB; AAD27078.
XX
PT Composition for modulating muscle cell and tissue contractility for
PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
PT comprising domains from urokinase-type plasminogen activator -

XX PS Claim 11; Fig 1D; 117pp; English.

CC The invention relates to a composition comprising one or more domains of

CC urokinase-type plasminogen activator (uPA). The composition is used to

CC modulate the contractility and angiogenic activity of a mammalian muscle,

CC endothelial cell or tissue. The composition is used for treating stroke,

CC hypotension, hypertension, atherosclerosis, heart attack, microvascular

CC occlusions, thrombotic microangiopathies, surgically induced thrombotic

CC disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell

CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,

CC diabetic retinopathy, wound healing, clotting disorder, uterine

CC contraction disorder, male impotence, respiratory disease or condition

CC such as asthma, adult respiratory distress syndrome, primary pulmonary

CC hypertension, microvascular thrombotic occlusion, and a disorder

CC associated with chronic intrapulmonary fibrin formation. The present

CC sequence is human urokinase-type plasminogen activator (uPA) amino

CC terminal fragment (ATF).

XX SQ Sequence 135 AA;

Query Match 99.0%; Score 494; DB 23; Length 135;

Best Local Similarity 98.8%; Pred. No. 2.7e-41;

Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGLGKHNYCRNPDN 60

DB 48 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGLGKHNYCRNPDN 107

QY 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86

DB 108 RRRPWCYVQVGLKPLVQECMVHDCAD 133

RESULT 4

AAE16549

ID AAE16549 standard; Protein; 143 AA.

XX AC AAE16549;

DT 09-APR-2002 (first entry)

DE Human uPA amino terminal fragment (ATF) and connecting peptide.

XX KW Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;

XX KW stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;

XX KW microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;

XX KW tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;

XX KW clotting disorder; uterine contraction disorder; respiratory disease;

XX KW adult respiratory distress syndrome; amino terminal fragment; ATF;

XX KW male impotence.

XX OS Homo sapiens.

XX PN WO200197752-A2.

XX PD 27-DEC-2001.

XX PF 13-JUN-2001; 2001WO-US18976.

XX PR 20-JUN-2000; 2000US-212874P.

XX PA (UYPE-) UNIV PENNSYLVANIA.

XX PI Cines DB, Higazi AA;

XX DR WPI; 2002-122240/16.

XX DR N-PSDB; AAD27082.

XX PT Composition for modulating muscle cell and tissue contractility for

XX PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,

XX PT comprising domains from urokinase-type plasminogen activator -

XX PS Claim 24; Fig 1H; 117pp; English.

CC The invention relates to a composition comprising one or more domains of

CC urokinase-type plasminogen activator (uPA). The composition is used to

CC modulate the contractility and angiogenic activity of a mammalian muscle,

CC endothelial cell or tissue. The composition is used for treating stroke,

CC hypotension, hypertension, atherosclerosis, heart attack, microvascular

CC occlusions, thrombotic microangiopathies, surgically induced thrombotic

CC disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell

CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,

CC diabetic retinopathy, wound healing, clotting disorder, uterine

CC contraction disorder, male impotence, respiratory disease or condition

CC such as asthma, adult respiratory distress syndrome, primary pulmonary

CC hypertension, microvascular thrombotic occlusion, and a disorder

CC associated with chronic intrapulmonary fibrin formation. The present

CC sequence is human urokinase-type plasminogen activator (uPA) amino

CC terminal fragment (ATF) and connecting peptide.

XX SQ Sequence 143 AA;

Query Match 99.0%; Score 494; DB 23; Length 143;

Best Local Similarity 98.8%; Pred. No. 2.8e-41;

Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGLGKHNYCRNPDN 60

DB 48 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGLGKHNYCRNPDN 107

QY 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86

DB 108 RRRPWCYVQVGLKPLVQECMVHDCAD 133

RESULT 5

AAG75492

ID AAG75492 standard; Protein; 337 AA.

XX AC AAG75492;

DT 03-SEP-2001 (first entry)

DE Human colon cancer antigen protein SEQ ID NO:6256.

XX KW Human; colon cancer; colon cancer antigen; diagnosis; detection;

XX KW colorectal carcinoma; chromosome 10.

XX OS Homo sapiens.

XX PN WO200122920-A2.

XX PD 05-APR-2001.

XX PF 28-SEP-2000; 2000WO-US26524.

XX PR 29-SEP-1999; 99US-0157137.

XX PR 03-NOV-1999; 99US-0163280.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Ruben SM, Barash SC, Birse CE, Rosen CA;

XX DR WPI; 2001-235357/24.

XX DR N-PSDB; AAH34897.

XX PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,

XX PT useful for preventing, diagnosing and/or treating colorectal cancers -

XX PS Claim 11; Page 7707-7708; 9803pp; English.

XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon

XX cancer-associated nucleic acid molecules (N) and proteins (P), where

XX the proteins are collectively known as colon cancer antigens. The colon

XX cancer antigens have cytostatic activity and can be used in gene

CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing P.
CC Inactive proteins or to supplement the patient's own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAB77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.

SO Sequence 337 AA;

Query Match 99.0%; Score 494; DB 22; Length 337;
Best Local Similarity 98.8%; Pred. No. 6.5e-41;
Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQGLGKHNYCRNPDN 60
DB 74 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQGLGKHNYCRNPDN 133

QY 61 RRRPWCYQVGLKPLVQECMVHDCAD 86
DB 134 RRRPWCYQVGLKPLVQECMVHDCAD 159

RESULT 6
ABP41795

ID ABP41795 standard; Protein; 337 AA.

XX AC ABP41795;

DT 22-AUG-2002 (first entry)

DE Human ovarian antigen HVVGB79, SEQ ID NO:2927.

XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
KW inflammatory condition; immune disorder; blood disorder;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disorder; urinary system disorder; drug screening;
KW gene therapy; chromosome mapping; forensic analysis;
KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW antiinflammatory; gynaecological; reproductive; chromosome 10q24.

XX Homo sapiens.

PN WO200200677-A1.

XX 03-JAN-2002.

XX 07-JUN-2001; 2001WO-US18569.

XX 07-JUN-2000; 2000US-209467P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Birse CE, Rosen CA;

XX WPI; 2002-147878/19.

XX N-PSDB; ABQ54872.

PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT useful in the prevention, treatment and diagnosis of cancer (e.g.
PT ovarian cancer), immune disorders, cardiovascular disorders and
PT neurological diseases -

XX Claim 11; SEQ ID NO 2927; 2922pp; English.

XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
XX ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
XX encompasses polypeptides 90% identical and polynucleotides 95% identical
XX to the sequences of the invention. The invention additionally relates to
XX recombinant vectors and host cells comprising human ovarian antigen
XX polynucleotides, antibodies against human ovarian antigens, and the use
XX of ovarian antigen polynucleotides and polypeptides in diagnosing,
XX treating, prognosing or preventing various ovary and/or breast-related
XX disorders. Such conditions include ovarian cancer and breast cancer, and
XX metastatic tumours of ovarian or breast origin, reproductive system
XX disorders (e.g., infertility, disorders of pregnancy, anovulation,
XX polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
XX disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
XX shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
XX vaginitis), immune disorders (e.g., congenital and acquired
XX immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
XX blood-related disorders (e.g., anaemia), cardiovascular disorders,
XX respiratory disorders, neurological disorders, gastrointestinal disorders
XX and urinary system disorders. Ovarian antigen polypeptides and
XX polynucleotides may also be used in screening for compounds which
XX modulate ovarian antigen expression or activity. The polynucleotides may
XX further be used for gene therapy, chromosome mapping, in the
XX identification of individuals and in forensic analysis, and the
XX polypeptides may be used as food additives or to prepare antibodies
XX useful in disease diagnosis, drug targeting and phenotyping. The present
XX sequence represents a human ovarian antigen of the invention.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 337 AA;

Query Match 99.0%; Score 494; DB 23; Length 337;
Best Local Similarity 98.8%; Pred. No. 6.5e-41;
Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQGLGKHNYCRNPDN 60
DB 74 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQGLGKHNYCRNPDN 133

QY 61 RRRPWCYQVGLKPLVQECMVHDCAD 86
DB 134 RRRPWCYQVGLKPLVQECMVHDCAD 159

RESULT 7
AAR68854

ID AAR68854 standard; protein; 365 AA.

XX AAR68854;

XX 22-NOV-1995 (first entry)

XX Delta 1-46 urokinase.

XX Human; des-epidermal growth factor homologous plasminogen activator;
KW uPA; liver membrane; reduced affinity; EGF homologous; thrombosis;
KW thrombolytic; increased half-life; urokinase.

XX Homo sapiens (engineered).

XX Key Location/Qualifiers

XX Misc-difference 1 /note= "amino acids 1-46 of wild-type urokinase
XX have been deleted"

XX US5376547-A.

XX 27-DEC-1994.

PF 30-JAN-1987; 87US-0008795.
XX
PR 30-JAN-1987; 87US-0008795.
PR 29-JAN-1988; 88US-0150267.
XX
PA (AMHP) AMERICAN HOME PROD CORP.
XX
PI Hung PP, Kalyan NK, Lee SL;
XX
DR WPI; 1995-043464/06.
XX
PT New modified plasminogen activator cpds. - having regions removed
PT to reduce affinity for liver membranes and increase circulation
PT half-life.
XX
PS Claim 1; ; 26pp; English.
XX
CC Amino acid residues 1-46 contain the EGF region of human urokinase.
CC Deletion of this region results in a plasminogen activator with
CC reduced affinity for liver cell membranes; the mutant protein is
CC not cleared from the circulation as rapidly as its wild-type tPA.
CC The specification only gives the sequence around the deletion and
CC not the full-length sequence of "delta 1-46 urokinase"; the
CC sequence in AAR68854 has been obtained by amending a previously
CC disclosed wild-type human urokinase sequence (from WO9501427)
CC according to the description given in Example 3.
XX
SQ Sequence 365 AA;

Query Match 99.0%; Score 494; DB 16; Length 365;
Best Local Similarity 98.8%; Pred. No. 7e-41;
Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGKHNCRNPDN 60
|||||
Db 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGKHNCRNPDN 61
|||||
OY 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
|||||
Db 62 RRRPWCYVQVGLKPLVQECMVHDCAD 87
|||||

RESULT 8
AAW13635
ID AAW13635 standard; Protein; 378 AA.

XX
AC AAW13635;

XX
DT 04-JUN-1997 (first entry)

XX
DE Human prourokinase variant lacking entire EGF domain.

XX
KW Human; prourokinase; hPUK; variant; half-life; increase; EGF;
KW epidermal growth factor domain; deletion; thrombolysis; fibrinolysis.

XX
OS Homo sapiens.
OS Synthetic.

XX
FH Key Location/Qualifiers

FT Region 1..9 /note= "residues 1-9 of native hPUK"

FT Region 10..378 /note= "residues 43-411 of native hPUK"

FT Misc-difference 144 /note= "corresponds to TAC codon"

XX
PN EP398361-A.

XX
PD 22-NOV-1990.

XX
PF 18-MAY-1990; 90EP-0109472.

XX
PR 22-FEB-1990; 90JP-0042020.

PR 18-MAY-1989; 89JP-0126433.
PR 03-JUL-1986; 86JP-0156936.
PR 18-FEB-1987; 87JP-0036495.
PR 18-MAY-1989; 89JP-0126434.

XX
PA (GREG) GREEN CROSS CORP.

XX
PI Airmura H, Amatsuji Y, Hirose M, Kasai S, Kawabe H;

PI Morita M, Tanabe T;

XX
DR WPI; 1990-350146/47.

DR N-PSDB; AAT61672.

XX
PT Human pro-urokinase variants - deficient in loop regions of
PT epidermal growth factor, showing long blood half-life, as
PT fibrinolytic agent

XX
PS Claim 1; Page -; 22pp; English.

XX
CC New variants of human prourokinase (hPUK) comprise a hPUK deficient
CC in (i) at least part of the first loop region of the epidermal growth
CC factor (EGF) domain; (ii) at least part of the first loop and at
CC least part of the second loop; or (iii) at least part of the third
CC loop. The hPUK variants show an increased blood half-life comparable
CC to that of the whole EGF domain-deficient hPUK variant and urokinase
CC while retaining the same properties as those of hPUK. They have
CC potent thrombolytic activity and very little tendency to cause
CC spontaneous bleeding. The present sequence represents a specific variant
CC of hPUK which lacks the entire EGF domain; the sequence does not
CC appear in the specification and has been created using the
CC wild-type hPUK sequence and the junction sequence after
CC deletion, both of which are given (in Fig 1 and in Fig 2(3),
CC respectively).

XX
SQ Sequence 378 AA;

Query Match 99.0%; Score 494; DB 11; Length 378;
Best Local Similarity 98.8%; Pred. No. 7.2e-41;
Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGKHNCRNPDN 60
|||||
Db 15 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGKHNCRNPDN 74
|||||

OY 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
|||||

Db 75 RRRPWCYVQVGLKPLVQECMVHDCAD 100
|||||

RESULT 9

AAAR66266
ID AAR66266 standard; protein; 386 AA.

XX
AC AAR66266;

XX
DT 17-AUG-1995 (first entry)

XX
DE Bifunctional urokinase variant M33.

XX
KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.

XX
OS Synthetic.

XX
FH Key Location/Qualifiers

FT Region 1..365 /label= M4

FT /note= "unglycosylated prourokinase(Ser47-Leu411)"

FT Disulfide-bond 4..85

FT Disulfide-bond 25..67

FT Disulfide-bond 56..80

FT Disulfide-bond 102..233

FT Disulfide-bond 143..159

FT	Disulfide-bond	151..222
FT	Disulfide-bond	247..316
FT	Disulfide-bond	279..295
FT	Disulfide-bond	306..334
FT	Region	365..366
FT	/label= X1	
FT	/note= "peptide bond"	
FT	366..386	
FT	/label= Y1	
PN	DE4323754-C.	
XX		
PD	01-DEC-1994.	
XX		
PF	15-JUL-1993;	93DE-4323754.
XX		
PR	15-JUL-1993;	93DE-4323754.
XX		
PA	(CHEF) GRUENENTHAL GMBH.	
PI	Heinzel-Wieland R, Saunders DJ, Schneider J, Steffens GJ;	
PI	Wendt S;	
DR	WPI; 1995-015191/03.	
XX		
PT	New bifunctional urokinase derivs and related plasmids - with	
PT	improved fibrinolytic and thrombin inhibiting activities, for	
PT	treating cardiac and cerebral infarct, pulmonary embolism, etc	
XX		
PS	Example 1; Page 11 and Fig 1; 34pp; German.	
XX		
CC	Bifunctional urokinase derivatives corresponding to the formula	
CC	M ₄ -X1-Y1 are claimed (see features table). Sequences AAR66244-R66266	
CC	are specific examples of such derivs. which have both improved	
CC	fibrinolytic and thrombin-inhibiting activities, compared to known	
CC	plasmaogen activators or thrombin inhibitors. The proteins are	
CC	useful as thrombolytic agents, e.g. for treatment of arterial	
CC	occlusions, deep vein thrombosis, cardiac and cerebral infarction	
CC	and pulmonary embolism.	
XX		
SQ	Sequence 386 AA;	
Query Match	99.0%; Score 494; DB 16; Length 386;	
Best Local Similarity	98.8%; Pred. No. 7.4e-41;	
Matches 85; Conservative	1; Mismatches 0; Indels 0; Gaps 0;	
QY	1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQQTYYHAHRSNALQLGLGKHNYCRNPDN 60	
Db	2 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQQTYYHAHRSDALQLGLGKHNYCRNPDN 61	
QY	61 RRRPWCYVQVGLKPLVQECMVHDCAD 86	
Db	62 RRRPWCYVQVGLKPLVQECMVHDCAD 87	
RESULT 10		
ID	AAW13636 standard; Protein; 389 AA.	
AC	AAW13636;	
DT	04-JUN-1997 (first entry)	
DE	Human prourokinase variant lacking EGF domain loops 1 and 2.	
KW	Human; prourokinase; hPUK; variant; half-life; increase; EGF;	
XX	epidermal growth factor domain; deletion; thrombolysis; fibrinolysis.	
OS	Homo sapiens.	
XX	Synthetic.	
Key	Location/Qualifiers	
FTT	Region 1..10	

FT	/note= "residues 1-10 of native hPUK"
FT	11..389
FT	/note= "residues 33-411 of native hPUK"
FT	155
FT	/note= "corresponds to TAC codon"
XX	
PN	EP398361-A.
XX	
PD	22-NOV-1990.
XX	
PF	18-MAY-1990; 90EP-0109472.
XX	
XX	22-FEB-1990; 90JP-0042020.
PR	18-MAY-1989; 89JP-0126433.
PR	03-JUL-1986; 86JP-0156936.
PR	18-FEB-1987; 87JP-0036495.
PR	18-MAY-1989; 89JP-0126434.
XX	
PA	(GREC) GREEN CROSS CORP.
XX	
PI	Airmura H, Amatsuji Y, Hirose M, Kasai S, Kawabe H;
PI	Morita M, Tanabe T;
XX	
DR	WPI; 1990-350146/47.
DR	N-PSDB; AAT61673.
PT	Human pro-urokinase variants - deficient in loop regions of
PT	epidermal growth factor, showing long blood half-life, as
PT	fibrinolytic agent
XX	
PS	Claim 6; Page -: 22pp; English.
XX	
CC	New variants of human prourokinase (hPUK) comprise a hPUK deficient
CC	in (i) at least part of the first loop region of the epidermal growth
CC	factor (EGF) domain; (ii) at least part of the first loop and at
CC	least part of the second loop; or (iii) at least part of the third
CC	loop. The hPUK variants show an increased blood half-life comparable
CC	to that of the whole EGF domain-deficient hPUK variant and urokinase
CC	while retaining the same properties as those of hPUK. They have
CC	potent thrombolytic activity and very little tendency to cause
CC	spontaneous bleeding. The present sequence represents a specific
CC	variant of hPUK which lacks loops 1 and 2 of the EGF domain; the
CC	sequence does not appear in the specification and has been created
CC	using the wild-type hPUK sequence and the junction sequence after
CC	deletion, both of which are given (in Fig 1 and on page 8,
CC	respectively).
XX	
SO	Sequence 389 AA;
	Query Match 99.0%; Score 494; DB 11; Length 389;
	Best Local Similarity 98.8%; Pred. No. 7.5e-41;
	Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY	1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQQT YHNRSNALQLGLGKHN YCRNPDN 60
Db	26 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQQT YHNRSDALQLGLGKHN YCRNPDN 85
QY	61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
Db	86 RRRPWCYVQVGLKPLVQECMVHDCAD 111
	RESULT 11
ID	AAR66245 standard; protein; 390 AA.
XX	
AC	AAR66245;
XX	
DT	17-AUG-1995 (first entry)
XX	
DE	Bifunctional urokinase variant M12.
XX	
KW	fibrinolysis;thrombin inhibition;thrombolytic; anti-thrombotic;

KW urokinase; variant; mutein.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 1..365
FT /label= M4
FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Disulfide-bond 366..371
FT Region /label= X1
FT Region 372..390
FT /label= Y1
XX
PN DE4323754-C.
XX
PD 01-DEC-1994.
XX
PF 15-JUL-1993; 93DE-4323754.
XX
PR 15-JUL-1993; 93DE-4323754.
XX
PA (CHEF) GRUENENTHAL GMBH.
XX
PI Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
PI Wnendt S;
XX
DR WPI; 1995-015191/03.
XX
XX
PT New bifunctional urokinase derivs and related plasmids - with
PT improved fibrinolytic and thrombin inhibiting activities, for
PT treating cardiac and cerebral infarct, pulmonary embolism, etc
XX
PS Example 1; Page 10 and Fig 1; 34pp; German.
XX
CC Bifunctional urokinase derivatives corresponding to the formula
CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
CC are specific examples of such derivs. which have both improved
CC fibrinolytic and thrombin-inhibiting activities, compared to known
CC plasminogen activators or thrombin inhibitors. The proteins are
CC useful as thrombolytic agents, e.g. for treatment of arterial
CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
CC and pulmonary embolism.
XX
SQ Sequence 390 AA;
Query Match 99.0%; Score 494; DB 16; Length 390;
Best Local Similarity 98.8%; Pred. No. 7.5e-41;
Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGKHNCRNPDN 60
DB 2 KTCYEGNGHFYRGASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGKHNCRNPDN 61
QY 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
DB 62 RRRPWCYVQVGLKPLVQECMVHDCAD 87
RESULT 12
AAR66247
ID AAR66247 standard; protein; 390 AA.
XX
AC AAR66247;
XX

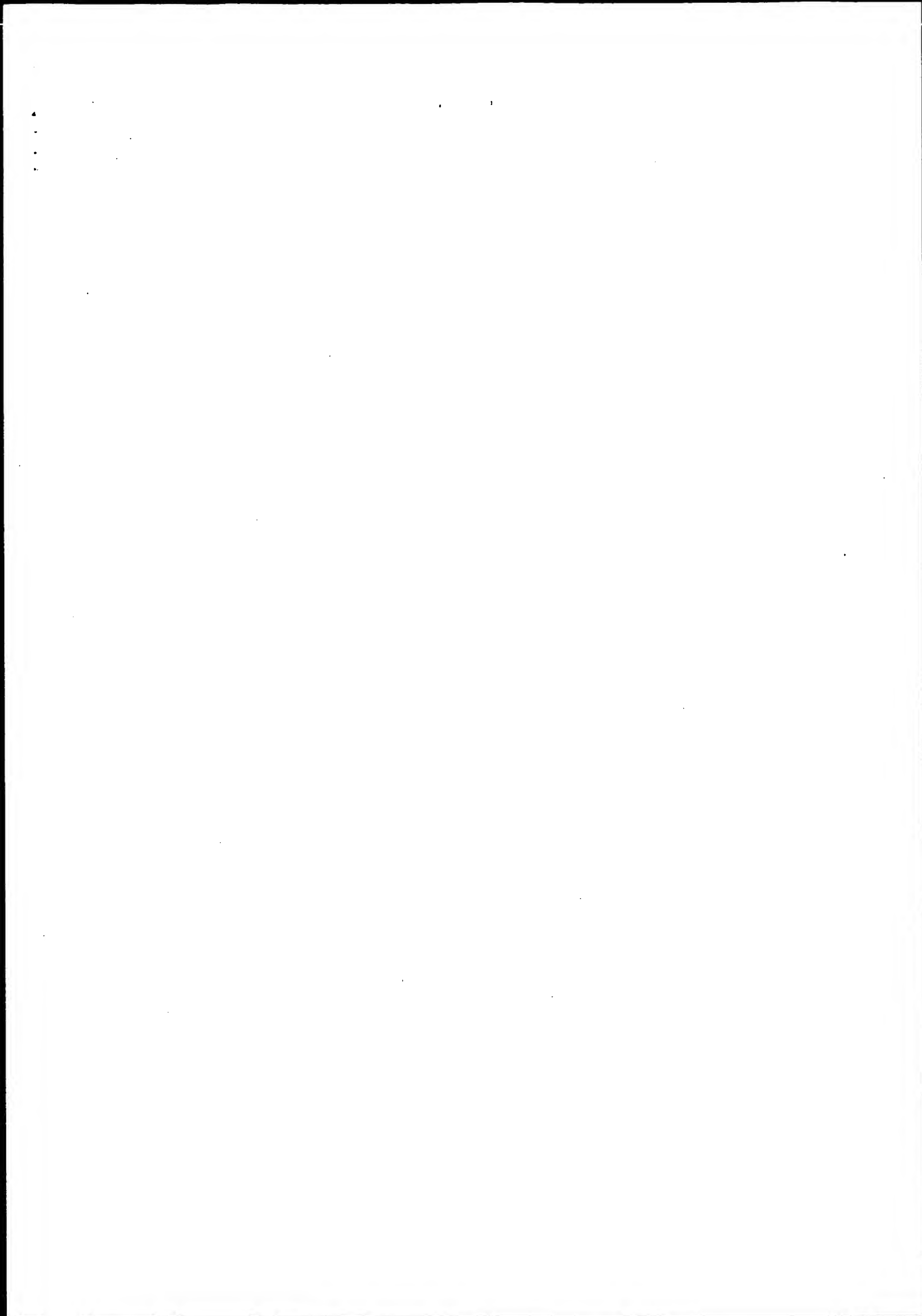
DT 17-AUG-1995 (first entry)
XX
DE Bifunctional urokinase variant M14.
XX
KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 1..365
FT /label= M4
FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Disulfide-bond 366..371
FT Region /label= X1
FT Region 372..390
FT /label= Y1
XX
PN DE4323754-C.
XX
PD 01-DEC-1994.
XX
PF 15-JUL-1993; 93DE-4323754.
XX
PR 15-JUL-1993; 93DE-4323754.
XX
PA (CHEF) GRUENENTHAL GMBH.
XX
PI Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
PI Wnendt S;
XX
DR WPI; 1995-015191/03.
XX
XX
PT New bifunctional urokinase derivs and related plasmids - with
PT improved fibrinolytic and thrombin inhibiting activities, for
PT treating cardiac and cerebral infarct, pulmonary embolism, etc
XX
PS Example 1; Page 10 and Fig 1; 34pp; German.
XX
CC Bifunctional urokinase derivatives corresponding to the formula
CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
CC are specific examples of such derivs. which have both improved
CC fibrinolytic and thrombin-inhibiting activities, compared to known
CC plasminogen activators or thrombin inhibitors. The proteins are
CC useful as thrombolytic agents, e.g. for treatment of arterial
CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
CC and pulmonary embolism.
XX
SQ Sequence 390 AA;
Query Match 99.0%; Score 494; DB 16; Length 390;
Best Local Similarity 98.8%; Pred. No. 7.5e-41;
Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGKHNCRNPDN 60
DB 2 KTCYEGNGHFYRGASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGKHNCRNPDN 61
QY 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
DB 62 RRRPWCYVQVGLKPLVQECMVHDCAD 87
RESULT 13

ID	AA	6254	standard; protein; 392 AA.
XX	AC	AA66254;	
XX	DT	17-AUG-1995	(first entry)
XX	DE	Bifunctional urokinase variant M21.	
XX	KW	fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic; urokinase; variant; mutein.	
OS	XX	Synthetic.	
FH	Key	Location/Qualifiers	
FT	Region	1..365	/label= M4
FT		/note= "unglycosylated prourokinase(Ser47-Leu411)"	
FT	Disulfide-bond	4..85	
FT	Disulfide-bond	25..67	
FT	Disulfide-bond	56..80	
FT	Disulfide-bond	102..233	
FT	Disulfide-bond	143..159	
FT	Disulfide-bond	151..222	
FT	Disulfide-bond	247..316	
FT	Disulfide-bond	279..295	
FT	Disulfide-bond	306..334	
FT	Disulfide-bond	366..371	
FT	Region	/label= X1	
FT		372..392	
FT		/label= Y1	
XX	DE	4323754-C.	
XX	PD	01-DEC-1994.	
XX	PE	15-JUL-1993;	93DE-4323754.
XX	PR	15-JUL-1993;	93DE-4323754.
XX	PA	(CHEF) GRUENENTHAL GMBH.	
XX	PI	Heinzel-Wieland R, Saunders DJ, Schneider J, Steffens GJ;	
XX	PI	Wendt S;	
DR	WP	1995-015191/03.	
XX	PT	New bifunctional urokinase derivs and related plasmids - with improved fibrinolytic and thrombin inhibiting activities, for treating cardiac and cerebral infarct, pulmonary embolism, etc	
XX	PS	Example 1; Page 10 and Fig 1; 34pp; German.	
XX	CC	Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266 are specific examples of such derivs. which have both improved fibrinolytic and thrombin-inhibiting activities, compared to known plasminogen activators or thrombin inhibitors. The proteins are useful as thrombolytic agents, e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac and cerebral infarction and pulmonary embolism.	
XX	SO	Sequence 392 AA;	
QY	Query Match	99.0%; Score 494; DB 16; Length 392;	
Db	Best Local Similarity	98.8%; Pred. No. 7.5e-41;	
Matches	85; Conservative	1; Mismatches 0; Indels 0; Gaps 0;	
QY	1	KTCEGNGHFYRGKASTDTMGRPCLPWN SATVLQOTYHNRSNALQLGKHNCRPN 60	
Db	2	KTCEGNGHFYRGKASTDTMGRPCLPWN SATVLQOTYHNRSDALQLGKHNCRPN 61	
QY	61	RRRPWCYVQVGKPLVQECMVHDCAD 86	

```

Db      62 RRRPWCYVQVGLKPLVQECMVHDCAD 87
|||||
RESULT 14
ID      AAR66255 standard; protein; 392 AA.
XX      AAR66255;
AC      17-AUG-1995 (first entry)
XX      DE
XX      Bifunctional urokinase variant M22.
KW      fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW      urokinase; variant; mutein.
XX      Synthetic.
OS      Key
XX      Location/Qualifiers
FH      Region
FT      1..365
FT      /label= M4
FT      /note= "unglycosylated prourokinase(Ser47-Leu411)"
FT      Disulfide-bond
FT      4..85
FT      Disulfide-bond
FT      25..67
FT      Disulfide-bond
FT      56..80
FT      Disulfide-bond
FT      102..233
FT      Disulfide-bond
FT      143..159
FT      Disulfide-bond
FT      151..222
FT      Disulfide-bond
FT      247..316
FT      Disulfide-bond
FT      279..295
FT      Disulfide-bond
FT      306..334
FT      Disulfide-bond
FT      366..371
FT      Region
FT      /label= X1
FT      372..392
FT      /label= Y1
XX      DE4323754-C.
XX      01-DEC-1994.
XX      15-JUL-1993; 93DE-4323754.
XX      PF
XX      15-JUL-1993; 93DE-4323754.
XX      PR
XX      (CHEF ) GRUENENTHAL GMBH.
XX      PA
XX      Heinzel-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
XX      Wnendt S;
XX      WPI; 1995-015191/03.
XX      DR
XX      New bifunctional urokinase derivs and related plasmids - with
XX      improved fibrinolytic and thrombin inhibiting activities, for
XX      treating cardiac and cerebral infarct, pulmonary embolism, etc
XX      Example 1; Page 10 and Fig 1; 34pp; German.
XX      PS
XX      Bifunctional urokinase derivatives corresponding to the formula
XX      M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
XX      are specific examples of such derivs. which have both improved
XX      fibrinolytic and thrombin-inhibiting activities, compared to known
XX      plasminogen activators or thrombin inhibitors. The proteins are
XX      useful as thrombolytic agents, e.g. for treatment of arterial
XX      occlusions, deep vein thrombosis, cardiac and cerebral infarction
XX      and pulmonary embolism.
XX      SQ
XX      Sequence 392 AA;
Query Match 99.0%; Score 494; DB 16; Length 392;
Best Local Similarity 98.8%; Pred. No. 7.5e-41;
Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2002, 12:58:41 ; Search time 14 Seconds

(without alignments)
88.571 Million cell updates/sec

Title: PCT-US02-27855-1

Perfect score: 499
Sequence: 1 KTCYEGNGHFYRGKASTDTM.....YVQVGLKPLVQECMVHDCAD 86Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 92612 seqs, 14418503 residues

Total number of hits satisfying chosen parameters: 92612

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*

```
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	494	99.0	88	US-09-880-503-1	Sequence 1, Appli
2	494	99.0	96	US-09-880-503-9	Sequence 9, Appli
3	494	99.0	135	US-09-880-503-4	Sequence 4, Appli
4	494	99.0	138	US-09-884-186-12	Sequence 12, Appli
5	494	99.0	143	US-09-880-503-8	Sequence 8, Appli
6	494	99.0	403	US-09-880-503-6	Sequence 6, Appli
7	494	99.0	411	US-09-880-503-3	Sequence 3, Appli
8	494	99.0	431	US-09-264-468B-1	Sequence 1, Appli
9	216	43.3	372	US-09-084-491A-3	Sequence 3, Appli
10	216	43.3	562	US-09-969-271-7	Sequence 7, Appli
11	215	43.1	562	US-09-974-298-145	Sequence 145, App
12	188.5	37.8	560	US-09-912-559-3	Sequence 3, Appli
13	188.5	37.8	560	US-09-912-559-4	Sequence 4, Appli
14	151	30.3	368	US-09-761-120-42	Sequence 42, Appli
15	149	29.9	378	US-09-873-676-1	Sequence 1, Appli
16	149	29.9	458	US-09-946-893-4	Sequence 4, Appli
17	149	29.9	569	US-09-946-893-5	Sequence 5, Appli
18	149	29.9	571	US-09-946-893-8	Sequence 8, Appli
19	149	29.9	576	US-09-946-893-6	Sequence 6, Appli

20	149	29.9	791	9	US-09-967-386-1	Sequence 1, Appli
21	149	29.9	810	10	US-09-946-893-2	Sequence 2, Appli
22	148.5	29.8	78	10	US-09-753-064-5	Sequence 5, Appli
23	148.5	29.8	78	10	US-09-761-120-23	Sequence 23, Appli
24	148.5	29.8	352	10	US-09-761-120-40	Sequence 40, Appli
25	147.5	29.5	937	9	US-09-974-298-129	Sequence 129, App
26	137	27.5	378	10	US-09-761-120-41	Sequence 41, Appli
27	137	27.5	459	10	US-09-761-120-46	Sequence 46, Appli
28	137	27.5	812	10	US-09-788-142-1	Sequence 1, Appli
29	137	27.5	812	10	US-09-761-120-1	Sequence 1, Appli
30	137	27.5	812	10	US-09-873-676-81	Sequence 81, Appli
31	136.5	27.4	339	10	US-09-788-142-3	Sequence 3, Appli
32	136.5	27.4	339	10	US-09-788-142-6	Sequence 6, Appli
33	136.5	27.4	339	10	US-09-761-120-3	Sequence 3, Appli
34	136.5	27.4	339	10	US-09-761-120-6	Sequence 6, Appli
35	135.5	27.2	78	10	US-09-761-120-22	Sequence 22, Appli
36	135.5	27.2	352	10	US-09-761-120-39	Sequence 39, Appli
37	134.5	27.0	339	10	US-09-788-142-5	Sequence 5, Appli
38	134.5	27.0	339	10	US-09-761-120-5	Sequence 5, Appli
39	133.5	26.8	339	10	US-09-788-142-4	Sequence 4, Appli
40	133.5	26.8	339	10	US-09-761-120-4	Sequence 4, Appli
41	132	26.5	72	10	US-09-848-288-6	Sequence 6, Appli
42	132	26.5	91	10	US-09-748-468-1	Sequence 1, Appli
43	132	26.5	263	10	US-09-084-491A-2	Sequence 2, Appli
44	132	26.5	339	10	US-09-788-142-2	Sequence 2, Appli
45	132	26.5	339	10	US-09-761-120-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-880-503-1
; Sequence 1, Application US/09880503

; Patent No. US20020131964A1

; GENERAL INFORMATION:

; APPLICANT: CINES, Douglas B

; APPLICANT: HIGAZI, Abd Al-Roof

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND

; FILE REFERENCE: 9596-331

; CURRENT APPLICATION NUMBER: US/09/880, 503

; CURRENT FILING DATE: 2001-06-13

; PRIOR APPLICATION NUMBER: US 60/212, 847

; PRIOR FILING DATE: 2000-06-20

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 88

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-09-880-503-1

Query Match 99.0%; Score 494; DB 10; Length 88;

Best Local Similarity 98.8%; Pred. No. 5.4e-51;

Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPMSATVLOQTYHAHRSNALDGLGKHNCRPN 60

Db 1 KTCYEGNGHFYRGKASTDTMGRPCLPMSATVLOQTYHAHRSNALDGLGKHNCRPN 60

QY 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86

Db 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86

RESULT 2
US-09-880-503-9

; Sequence 9, Application US/09880503

; Patent No. US20020131964A1

; GENERAL INFORMATION:

; APPLICANT: CINES, Douglas B

; APPLICANT: HIGAZI, Abd Al-Roof

;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
;; FILE REFERENCE: 9596-331
;; CURRENT APPLICATION NUMBER: US/09/880,503
;; PRIOR FILING DATE: 2001-06-13
;; PRIOR APPLICATION NUMBER: US 60/212,847
;; NUMBER OF SEQ ID NOS: 18
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO: 9
;; LENGTH: 96
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-880-503-9

Query Match 99.0%; Score 494; DB 10; Length 96;
Best Local Similarity 98.8%; Pred. No. 5.9e-51;
Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNYCRNPDN 60
|||
DB 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNYCRNPDN 60
|||
OY 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
|||
DB 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86

RESULT 3
US-09-880-503-4
; Sequence 4, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 4
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-4

Query Match 99.0%; Score 494; DB 10; Length 135;
Best Local Similarity 98.8%; Pred. No. 8.4e-51;
Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNYCRNPDN 60
|||
DB 48 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNYCRNPDN 107
|||
OY 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
|||
DB 108 RRRPWCYVQVGLKPLVQECMVHDCAD 133

RESULT 4
US-09-984-186-12
; Sequence 12, Application US/09984186
; Patent No. US20020151011A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guitton, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice

;; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
;; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
;; CONTAINING SAID POLYPEPTIDES
;; NUMBER OF SEQUENCES: 36
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Rhone-Poulenc Rorer Inc.
;; STREET: 500 Arcola Road, 3C43
;; CITY: Collegeville
;; STATE: PA
;; COUNTRY: USA
;; ZIP: 19426
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: Macintosh
;; OPERATING SYSTEM: System 7.1
;; SOFTWARE: Word 5.1 (Patentin)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/984,186
;; FILING DATE: 29-Oct-2001
;; CLASSIFICATION: <unknown>
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/797,689
;; FILING DATE: 31-JAN-1997
;; APPLICATION NUMBER: US 08/256,927
;; FILING DATE: 28-JUL-1994
;; APPLICATION NUMBER: FR 92/01064
;; FILING DATE: 31-JAN-1992
;; APPLICATION NUMBER: PCT/FR93/00085
;; FILING DATE: 28-JAN-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Smith Ph.D., Julie K.
;; REGISTRATION NUMBER: P-38,619
;; REFERENCE/DOCKET NUMBER: ST92006-US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (610) 454-3839
;; TELEFAX: (610) 454-3808
;; INFORMATION FOR SEQ ID NO: 12:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 138 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-984-186-12

Query Match 99.0%; Score 494; DB 10; Length 138;
Best Local Similarity 98.8%; Pred. No. 8.6e-51;
Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNYCRNPDN 60
|||
DB 51 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNYCRNPDN 110
|||
OY 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
|||
DB 111 RRRPWCYVQVGLKPLVQECMVHDCAD 136

RESULT 5
US-09-880-503-8
; Sequence 8, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-8

Query Match 99.0%; Score 494; DB 10; Length 143;
Best Local Similarity 98.8%; Pred. No. 9e-51;
Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGLGKHNCRPN 60
DB 48 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGLGKHNCRPN 107
QY 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
DB 108 RRRPWCYVQVGLKPLVQECMVHDCAD 133

RESULT 6
US-09-880-503-6

; Sequence 6, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:

; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Rooof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-6

Query Match 99.0%; Score 494; DB 10; Length 403;
Best Local Similarity 98.8%; Pred. No. 2.7e-50;
Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGLGKHNCRPN 60
DB 48 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGLGKHNCRPN 107
QY 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
DB 108 RRRPWCYVQVGLKPLVQECMVHDCAD 133

RESULT 7
US-09-880-503-3

; Sequence 3, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:

; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Rooof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 411

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-3

Query Match 99.0%; Score 494; DB 10; Length 411;
Best Local Similarity 98.8%; Pred. No. 2.7e-50;
Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGLGKHNCRPN 60
DB 48 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGLGKHNCRPN 107
QY 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
DB 108 RRRPWCYVQVGLKPLVQECMVHDCAD 133

RESULT 8
US-09-264-468B-1

; Sequence 1, Application US/09264468B
; Patent No. US20020106775A1
; GENERAL INFORMATION:

; APPLICANT: Wang, Jieyi
; APPLICANT: Nienaber, Vicki L.
; APPLICANT: Henkin, Jack
; APPLICANT: Smith, Richard A.
; APPLICANT: Walter, Karl A.
; APPLICANT: Severin, Jean M.
; APPLICANT: Edalji, Rohinton
; APPLICANT: Johnson Jr., Robert W.
; APPLICANT: Holzman, Thomas F.
; TITLE OF INVENTION: HIGHLY CRYSTALLINE UROKINASE
; FILE REFERENCE: 6310.US.PI
; CURRENT APPLICATION NUMBER: US/09/264,468B
; CURRENT FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: US 09/036,361
; PRIOR FILING DATE: 1998-03-06
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 431

; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(20)
; OTHER INFORMATION: Leader sequence
; NAME/KEY: VARIANT
; LOCATION: (279)...(279)
; OTHER INFORMATION: Xaa = any amino acid
; NAME/KEY: VARIANT
; LOCATION: (302)...(302)
; OTHER INFORMATION: Xaa = any amino acid
US-09-264-468B-1

Query Match 99.0%; Score 494; DB 10; Length 431;
Best Local Similarity 98.8%; Pred. No. 2.9e-50;
Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGLGKHNCRPN 60
DB 68 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGLGKHNCRPN 127
QY 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
DB 128 RRRPWCYVQVGLKPLVQECMVHDCAD 153

RESULT 9
US-09-084-491A-3

; Sequence 3, Application US/09084491A
; Patent No. US20020061576A1
; GENERAL INFORMATION:

APPLICANT: MOORE, PAUL A.
APPLICANT: RUBEN, STEVEN M.
APPLICANT: EBNER, REINHARD
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE PROTEASE
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/084,491A
FILING DATE: 27-MAY-1998
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: BROOKES, ANDERS A.
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF378
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8439

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 372 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-084-491A-3

Query Match 43.3%; Score 216; DB 10; Length 372;
Best Local Similarity 47.0%; Pred. No. 6.2e-18;
Matches 39; Conservative 9; Mismatches 35; Indels 0; Gaps 0;

QY 3 CYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGLGKHNYCRNPDNR 62
DB 25 CYFGNSAYRGTHSLTESGASCLPWNMSMILIGKYTAQNPSAQLGLGKHNYCRNPDGDA 84
QY 63 RPMCYYOVGLKPLVQECMVHDC 85
DB 85 KPWCHVLKNRRLTWEXCDVPSCS 107

RESULT 10
US-09-969-271-7

Sequence 7, Application US/09969271
Patent No. US20020098179A1
GENERAL INFORMATION:
APPLICANT: Pfizer Inc. (All designated states except GB and EP (GB));
APPLICANT: Pfizer Limited (GB and EP (GB) only)
TITLE OF INVENTION: Pharmaceutical Combinations
FILE REFERENCE: PCS10951APME
CURRENT APPLICATION NUMBER: US/09/969,271
CURRENT FILING DATE: 2001-10-01
PRIOR APPLICATION NUMBER: GB 0025473.0
PRIOR FILING DATE: 2000-10-17
NUMBER OF SEQ ID NOS: 7
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 562
TYPE: PRT
ORGANISM: Homo sapiens
US-09-969-271-7

Query Match 43.3%; Score 216; DB 10; Length 562;
Best Local Similarity 47.0%; Pred. No. 9.7e-18;

Matches 39; Conservative 9; Mismatches 35; Indels 0; Gaps 0;
QY 3 CYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGLGKHNYCRNPDNR 62
DB 215 CYFGNSAYRGTHSLTESGASCLPWNMSMILIGKYTAQNPSAQLGLGKHNYCRNPDGDA 274
QY 63 RPMCYYOVGLKPLVQECMVHDC 85
DB 275 KPWCHVLKNRRLTWEXCDVPSCS 297

RESULT 11
US-09-974-298-145

Sequence 145, Application US/09974298
Patent No. US20020156263A1
GENERAL INFORMATION:
APPLICANT: Chen, Huel-Mei
TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
FILE REFERENCE: PA-0037 P
CURRENT APPLICATION NUMBER: US/09/974,298
PRIOR APPLICATION NUMBER: 60/238,331
PRIOR FILING DATE: 2000-05-10
NUMBER OF SEQ ID NOS: 194
SOFTWARE: PERL Program
SEQ ID NO 145
LENGTH: 562
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No. US20020156263A1 1001470CD1
US-09-974-298-145

Query Match 43.1%; Score 215; DB 9; Length 562;
Best Local Similarity 45.9%; Pred. No. 1.3e-17;
Matches 39; Conservative 10; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGLGKHNYCRNPDNR 61
DB 126 TCYEDOGISYRGTHSLTESGASCLPWNMSMILIGKYTAQNPSAQLGLGKHNYCRNPDGDA 185
QY 62 RPMCYYOVGLKPLVQECMVHDCAD 86
DB 186 SKPWCVYFKAGKYSSEFCSTPACSE 210

RESULT 12
US-09-912-559-3

Sequence 3, Application US/09912559
Patent No. US20020142316A1
GENERAL INFORMATION:
APPLICANT: ROEMISCH, JUERGEN
APPLICANT: STOEHR, HANS-ARNOLD
APPLICANT: FEUSSNER, ANNETTE
APPLICANT: LANG, WIEGAND
APPLICANT: WEIMER, THOMAS
APPLICANT: BECKER, MARGRET
APPLICANT: NERLICH, CLAUDIA
APPLICANT: MUTH-NAUMANN, GUDRUN
TITLE OF INVENTION: MUTANTS OF THE FACTOR VII-ACTIVATING PROTEASE AND
DETECTION METHODS USING SPECIFIC ANTIBODIES
FILE REFERENCE: 06478.1457
CURRENT APPLICATION NUMBER: US/09/912,559
CURRENT FILING DATE: 2001-07-26
PRIOR APPLICATION NUMBER: DE 100 36 641.4
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: DE 100 50 040.4
PRIOR FILING DATE: 2000-10-10
PRIOR APPLICATION NUMBER: DE 100 52 319.6
PRIOR FILING DATE: 2000-10-21
PRIOR APPLICATION NUMBER: DE 101 18 706.8
PRIOR FILING DATE: 2001-04-12

100

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2002, 12:57:46 ; Search time 38 Seconds
(without alignments)
217.567 Million cell updates/sec

Title: PCT-US02-27855-1
Perfect score: 499
Sequence: 1 KTCYEGNGHRYRGKASTDTM.....YVQVGLKPLVQECMVHDCAD 86

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	494	99.0	431	1	UKHU u-plasminogen acti
2	459	92.0	433	1	UKBAY u-plasminogen acti
3	409	82.0	442	1	UKPG u-plasminogen acti
4	400	80.2	432	1	S18932 u-plasminogen acti
5	383	76.8	433	1	UKN560 u-plasminogen acti
6	381	76.4	433	1	UKMS u-plasminogen acti
7	236	47.3	431	2	JS0599 t-plasminogen acti
8	236	47.3	477	1	A34369 t-plasminogen acti
9	236	47.3	477	1	JS0598 t-plasminogen acti
10	226	45.3	559	1	A35029 t-plasminogen acti
11	218	43.7	394	2	JS0600 t-plasminogen acti
12	216	43.3	562	1	UKHUT t-plasminogen acti
13	215	43.1	291	2	I38098 t-plasminogen acti
14	215	43.1	559	1	A29941 t-plasminogen acti
15	208	41.7	477	2	JS0597 t-plasminogen acti
16	204.5	41.0	558	2	JS0597 plasma hyaluronan-
17	203.5	40.8	434	1	A35005 u-plasminogen acti
18	193	38.7	655	1	A46688 hepatocyte growth
19	188.5	37.8	560	1	JC4795 plasma hyaluronan-
20	179.5	36.0	603	2	S28941 coagulation factor
21	170.5	34.2	615	1	KEHU12 coagulation factor
22	158	31.7	89	2	A60140 plasmin (EC 3.4.21
23	156	31.3	120	2	E61545 apolipoprotein(a)
24	156	31.3	1420	2	A32869 coagulation factor
25	155.5	31.2	593	2	S45281 coagulation factor
26	155	31.1	4548	1	S00657 apoprotein(a) (EC
27	154	30.9	812	1	PLBO plasmin (EC 3.4.21
28	153	30.7	790	1	PLPG plasmin (EC 3.4.21
29	152.5	30.6	169	2	A40522 plasmin (EC 3.4.21

30	150	30.1	123	2	C61545 plasmin (EC 3.4.21
31	149	29.9	810	1	PLHU plasmin (EC 3.4.21
32	147.5	29.6	460	2	B61545 plasmin (EC 3.4.21
33	147.5	29.6	810	2	I46260 plasmin (EC 3.4.21
34	147	29.5	937	2	A45082 neurotrophic recep
35	146	29.3	810	2	B30848 plasmin (EC 3.4.21
36	143.5	28.8	711	1	A47136 macrophage-stimula
37	143.5	28.8	943	2	B45082 neurotrophic recep
38	142	28.5	806	2	T18840 plasmin (EC 3.4.21
39	137	27.5	812	1	PLMS plasmin (EC 3.4.21
40	134	26.9	455	2	A61545 hepatocyte growth
41	134	26.9	685	1	A48289 hepatocyte growth
42	133.5	26.8	728	1	A60185 hepatocyte growth
43	131.5	26.4	411	2	I51285 macrophage-stimula
44	131.5	26.4	716	1	A40332 macrophage-stimula
45	131.5	26.4	728	1	JH0579 hepatocyte growth

ALIGNMENTS

RESULT 1

UKHU

u-plasminogen activator (EC 3.4.21.73) precursor [validated] - human
N;Alternate names: cellular plasminogen activator; urokinase; urokinase-type plasminogen
N;Contents: urokinase-type plasminogen activator chain A; urokinase-type plasminogen
in form

C;Species: Homo sapiens (man)

C;Date: 17-Dec-1982 #sequence_revision 04-Dec-1986 #text_change 15-Sep-2000

C;Accession: A00931; I52209; JT0102; A37561; I38102; S65783; A37562; A37563; A37564;

R;Riccio, A.; Grimaldi, G.; Verde, P.; Sebastio, G.; Boast, S.; Blasl, F.

Nucleic Acids Res. 13, 2759-2771, 1985

A;Title: The human urokinase-plasminogen activator gene and its promoter.

A;Reference number: A00931; MUID:85215647; PMID:2987867

A;Accession: A00931

A;Molecule type: DNA

A;Residues: 1-431 <RIC>

A;Cross-references: GB:X02419; NID:g37601; PIDN:CAA26268.1; PID:g1834524

A;Note: the authors translated the codon ATG for residue 214 as Ile

R;Nagamine, Y.; Pearson, D.; Gratlan, M.

Biochem. Biophys. Res. Commun. 132, 563-569, 1985

A;Title: Exon-intron boundary sliding in the generation of two mRNAs coding for porci

A;Reference number: I52209; MUID:86050639; PMID:3933505

A;Accession: I52209

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 145-161 <NAG1>

A;Cross-references: GB:K03027; NID:g340174; PIDN:AAA61257.1; PID:g340175

R;Nagai, M.; Hiramatsu, R.; Kaneda, T.; Hayasuke, N.; Arimura, H.; Nishida, M.; Suyam

Gene 36, 183-188, 1985

A;Title: Molecular cloning of cDNA coding for human preprourokinase.

A;Reference number: JT0102; MUID:86056954; PMID:2415429

A;Accession: JT0102

A;Molecule type: mRNA

A;Residues: 1-213, 'I', 215-431 <NAG2>

A;Cross-references: GB:K03226; NID:g340155; PIDN:AAC97138.1; PID:g340158; GB:D00244;

R;Jacobs, P.; Cravador, A.; Loriau, R.; Brockly, F.; Colau, B.; Chuchana, P.; van Els

Proc. Natl. Acad. Sci. U.S.A. 81, 4727-4731, 1984

A;Title: Identification and primary sequence of an unspliced human urokinase poly(A)+

A;Reference number: A37561; MUID:84272706; PMID:6589620

A;Accession: A37561

A;Molecule type: mRNA

A;Residues: 66-431 <VER>

A;Cross-references: GB:D00244; NID:g220138

R;Jacobs, P.; Cravador, A.; Loriau, R.; Brockly, F.; Colau, B.; Chuchana, P.; van Els

DNA 4, 139-146, 1985

A;Title: Molecular cloning, sequencing, and expression in Escherichia coli of human p

A;Reference number: I38102; MUID:85203359; PMID:3888571

A;Accession: I38102

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-150, 'W', 152-213, 'I', 215-385, 'C', 387-429, 'V', 431 <JAC>

A;Cross-references: EMBL:X02760; NID:g35297; PIDN:CAA26535.1; PID:g35298

R;Yoshimoto, M.; Ushiyama, Y.; Sakai, M.; Tamaki, S.; Hara, H.; Takahashi, K.; Sawasaki, K.; Blochim. Biophys. Acta 1293, 83-89, 1996
A:Title: Characterization of single chain urokinase-type plasminogen activator with a novel
A:Reference number: S65783; MUID:96186279; PMID:8652631
A:Accession: S65783
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 21-140, 'L', 142-213, 'I', 215-431 <YOS>
A:Cross-references: EMBL:D11143; NID:91311467; PIDN:BA01919.1; PID:91199928
R;Gunzler, W.A.; Steffens, G.J.; Otting, F.; Kim, S.M.A.; Frankus, E.; Flohe, L.
Hoppe-Seyler's Z. Physiol. Chem. 363, 1155-1165, 1982
A:Title: The primary structure of high molecular mass urokinase from human urine.
A:Reference number: A37562; MUID:83055084; PMID:6754569
A:Accession: A37562
A:Molecule type: protein
A:Residues: 21-177 <GUN>
R;Schaller, J.; Nick, H.; Rickl, E.E.; Gillesen, D.; Lergier, W.; Studer, R.O.
Eur. J. Biochem. 125, 251-257, 1982
A:Title: Human low-molecular-weight urinary urokinase. Partial characterization and pre
A:Reference number: A37563; MUID:83003608; PMID:6749491
A:Accession: A37563
A:Molecule type: protein
A:Residues: 156-176, 179-193, 'T', 195, 'T', 197-224 <SCH>
R;Steffens, G.J.; Gunzler, W.A.; Otting, F.; Frankus, E.; Flohe, L.
Hoppe-Seyler's Z. Physiol. Chem. 363, 1043-1058, 1982
A:Title: The complete amino acid sequence of low molecular mass urokinase from human uri
A:Reference number: A37564; MUID:83055099; PMID:6754572
A:Accession: A37564
A:Molecule type: protein
A:Residues: 158-410 <STE>
R;Kentzer, E.J.; Buko, A.; Menon, G.; Sarin, V.R.
Biochem. Biophys. Res. Commun. 171, 401-406, 1990
A:Title: Carbohydrate composition and presence of a fucose-protein linkage in recombinant
A:Reference number: A35689; MUID:90365737; PMID:2393398
A:Accession: A35689
A:Molecule type: protein
A:Residues: 21-30, 'X', 32, 'X', 34-38, 'X', 40-43 <KEN>
A:Note: Identification of a fucose and attempt to determine its attachment site
R;Rabban, S.A.; Desjardins, J.; Bell, A.W.; Banville, D.; Mazar, A.; Henkin, J.; Goltz, B.
Biochem. Biophys. Res. Commun. 173, 1058-1064, 1990
A:Title: An amino-terminal fragment of urokinase isolated from a prostate cancer cell li
A:Reference number: A36697; MUID:91097529; PMID:2125213
A:Accession: A36697
A:Molecule type: protein
A:Residues: 21-34 <RAB>
R;Li, X.; Bokman, A.M.; Llinas, M.; Smith, R.A.G.; Dobson, C.M.
submitted to the Brookhaven Protein Data Bank, July 1993
A:Reference number: A51255; PDB:1KDU
A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue
R;Li, X.; Smith, R.A.G.; Dobson, C.M.
Biochemistry 31, 9562-9571, 1992
A:Title: Sequential (1)H NMR assignments and secondary structure of the kringle domain f
A:Reference number: A44375; MUID:93003110; PMID:1327118
A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR
R;Hansen, A.P.; Petros, A.M.; Meadows, R.P.; Nettlesheim, D.G.; Mazar, A.P.; Olejniczak, Z.
submitted to the Brookhaven Protein Data Bank, January 1994
A:Reference number: A66822; PDB:1URK
A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue
R;Spraggon, G.S.; Phillips, C.; Nowak, U.K.; Ponting, C.P.; Saunders, D.; Dobson, C.M.;
submitted to the Brookhaven Protein Data Bank, July 1995
A:Reference number: A66058; PDB:1LMW
A:Contents: annotation; X-ray crystallography, 2.5 angstroms, residues 168-175; 179-426
C:Comment: This enzyme is found in urine in a high molecular mass form, consisting of A
C:Comment: Urokinase-type plasminogen activator proteolytically activates plasminogen, a
C:Genetics:
A:Gene: GDB:PLAU
A:Cross-references: GDB:119497; OMIM:191840
A:Map position: 10q24-10q24
A:Introns: 19/3; 29/1; 65/1; 123/2; 154/1; 227/2; 277/1; 324/1; 373/3
C:Function:
A:Description: proteolytically activates plasminogen
A:Pathway: fibrinolysis
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try

C:Keywords: fibrinolysis; glycoprotein; heterodimer; hydrolase; kringle; serine prote
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-431/Product: urokinase-type plasminogen activator, single chain form #status pre
F:21-177/Product: urokinase-type plasminogen activator chain A #status experimental <
F:31-62/Domain: EGF homology <EGF>
F:70-151/Domain: kringle homology <KR>
F:156-177/Product: urokinase-type plasminogen activator chain A1 #status experimental
F:179-431/Product: urokinase-type plasminogen activator chain B #status experimental
F:179-419/Domain: trypsin homology <TRY>
F:31-39, 33-51, 53-62, 70-151, 91-133, 122-146, 168-299, 209-225, 217-288, 313-382, 345-361, 372
F:38/Binding site: carbohydrate (Thr) (covalent) #status predicted
F:178-179/Cleavage site: Lys-Ile (plasmin) #status experimental
F:224, 275, 376/Active site: His, Asp, Ser #status experimental
F:322/Binding site: carbohydrate (Asn) (covalent) #status experimental
Query Match 99.0%; Score 494; DB 1; Length 431;
Best Local Similarity 98.8%; Pred. No. 4e-47;
Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGKHNCRPN 60
Db 68 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGKHNCRPN 127
QY 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
Db 128 RRRPWCYVQVGLKPLVQECMVHDCAD 153

RESULT 2

UKBAY

u-plasminogen activator (EC 3.4.21.73) precursor - yellow baboon
C:Species: Papio cynocephalus, Papio hamadryas cynocephalus (yellow baboon)
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Jun-1999
C:Accession: S14687; S08651
R;Au, Y.P.T.; Wang, T.W.; Clowes, A.W.
Nucleic Acids Res. 18, 3411, 1990
A:Title: Nucleotide and deduced amino acid sequences of baboon urokinase-type plasmin
A:Reference number: S14687; MUID:90287734; PMID:2113276
A:Accession: S14687
A:Molecule type: mRNA
A:Residues: 1-433 <AU>
A:Cross-references: EMBL:X51935; NID:938130; PIDN:CAA36200.1; PID:938131
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology;
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine protease
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-176/Product: plasminogen activator chain A #status predicted <ACH>
F:30-61/Domain: EGF homology <EGF>
F:69-150/Domain: kringle homology <KR>
F:178-433/Product: plasminogen activator chain B #status predicted <BCH>
F:178-421/Domain: trypsin homology <TRY>
F:167-298, 208-224, 216-287, 315-384, 347-363, 374-402/Disulfide bonds: #status predicted
F:223, 274, 378/Active site: His, Asp, Ser #status predicted
F:324/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 92.0%; Score 459; DB 1; Length 433;
Best Local Similarity 93.0%; Pred. No. 3.1e-43;
Matches 80; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGKHNCRPN 60
Db 67 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGKHNCRPN 126
QY 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
Db 127 RRRPWCYVQVGLKPLVQECMVHDCAD 152

RESULT 3

UKPG

u-plasminogen activator (EC 3.4.21.73) precursor - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 04-Dec-1986 #sequence_revision 17-Mar-1987 #text_change 07-Aug-1998

C;Accession: A00932
R;Nagamine, Y.; Pearson, D.; Altus, M.S.; Reich, E.
Nucleic Acids Res. 12, 9525-9541, 1984
A;Title: cDNA and gene nucleotide sequence of porcine plasminogen activator.
A;Reference number: A00932; MUID:85087954; PMID:6096832
A;Accession: A00932
A;Molecule type: DNA
A;Residues: 1-240,'H','242-442 <NAG1>
A;Experimental source: kidney cell line LLC-PK1
R;Nagamine, Y.
submitted to the Protein Sequence Database, December 1986
A;Reference number: A37566
A;Contents: annotation; correction to residue 241
C;Genetics:
A;Intons: 19/3; 31/1; 67/1; 125/2; 165/1; 238/2; 288/1; 335/1; 384/3
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-188/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F;33-64/Domain: EGF homology <EGF>
F;72-153/Domain: kringle homology <KR>
F;190-442/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
F;190-430/Domain: trypsin homology <TRY>
F;152/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;179-310-220-236,228-299,324-393,356-372,383-411/Disulfide bonds: #status predicted
F;235,286,387/Active site: His, Asp, Ser #status predicted

```
Query Match      82.0%; Score 409; DB 1; Length 442;
Best Local Similarity 82.4%; Pred. No. 1.le-37;
Matches 70; Conservative 8; Mismatches 7; Indels 0; Gaps 0;
```

RESULT 4
S18932
u-plasminogen activator (EC 3.4.21.73) precursor - rat
N;Alternate names: plasminogen activator, urokinase-type; urinary plasminogen activator
C;Species: Rattus norvegicus (Norway rat)
C;Date: 18-Oct-1989 #sequence_revision 10-Feb-1995 #text_change 18-Jun-1999
C;Accession: S24604; I60186; I53472; S18932

submitted to the EMBL Data Library, April 1992
A; Reference number: S24604

A/Accession: S24604
A/Molecule type: mRNA
A/Residues: 1-15, 'H', 17-23, 'G', 25-331, 'N', 333-432 <RAB>
A/Cross-references: EMBL:X65651; NID:g57456; PIDN:CAA46601.1; PID:g57457
A/Experimental source: tissue kidney
R/Henderson, B.R.; Tansey, W.P.; Phillips, S.M.; Ramshaw, I.A.; Keeford, R.F.
Cancer Res. 52, 2489-2496, 1992
A/Title: Transcriptional and posttranscriptional activation of urokinase plasminogen acti
A/Reference number: I60186; MUID:92233409; PMID:1568219

A;Accession: I60186
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA

A;Residues: 1-432 <RES>
A;Cross-references: EMBL:X63434; NID:g57465; PIDN:CAA45028.1; PID:g57466
A;Experimental source: strain Fischer 344; tissue mammary

R; Ragnano, P.; Cassano, S.; Degen, J.; Kessler, C.; Blasl, F.; Rossi, G.
FEBS Lett. 306, 193-198, 1992

A;Title: The receptor for the plasminogen activator of urokinase type is up-regulated in
A;Reference number: I53472; MUID:923339549; PMID:1321734
A;Accession: I53472

A:Accession: J33472
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

A;Residues: 31-62 <RE2>

```

A:Cross-references: EMBL:X66907; NID:g396200; PIDN:CAA47356.1; PID:g938279
C:Genetics:
A:Gene: uPA
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology;
C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-177/Product: urokinase-type plasminogen activator chain A #status predicted <ACH
F;31-62/Domain: EGF homology <EGF>
F;70-151/Domain: kringle homology <KR>
F;179-432/Product: urokinase-type plasminogen activator chain B #status predicted <BC
F;179-420/Domain: trypsin homology <TRY>
F;168-300,210-226,218-289,314-383,346-362,373-401/Disulfide bonds: #status predicted
F;225,276,377/Active site: His, Asp, Ser #status predicted

Query Match      80.2%; Score 400; DB 1; Length 432;
Best Local Similarity 80.0%; Pred. No. 1.1e-36;
Matches 68; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

```

[illegible]

RESULT 5
JN0560
u-plasminogen activator (EC 3.4.21.73) precursor - bovine
N;Alternate names: uPA
C;Species: Bos primigenius taurus (cattle)
C;Date: 31-Dec-1993 #sequence-revision 31-Dec-1993 #text-change 16-Jul-1999
C;Accession: JN0560
R;Kraetzschmar, J.; Haendler, B.; Kojima, S.; Rifkin, D.B.; Schleuning, W.D

Gene 125, 177-183, 1993
A;Title: Bovine urokinase-type plasminogen activator and its receptor: cloning and in
A;Reference number: JN0560; MUID:93216119; PMID:8385052

A;Accession: JN0560
A;Molecule type: mRNA
A;Residues: 1-433 <KRA>
A;Cross-references: GB:L03546; NID:g163800; PIDN:AAA51419.1; PID:g163801
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology;
C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F;1-20/Domains: signal sequence #status predicted <SIG>
F;21-179/Product: plasminogen activator chain A #status predicted <MA1>
F;21-179/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>

F;33-64/Domain: EGF homology <EGF>
F;72-153/Domain: kringle homology <KR>
F;181-433/Product: plasminogen activator chain B #status predicted <MA2>

F;181-421/Domain: trypsin homology <TRY>
F;170-301,211-227,219-290,315-384,347-363,374-402/Disulfide bonds: #status predicted
F;226,277,378/Active site: His, Asp, Ser #status predicted

Query Match	76.8%;	Score 383;	DB 1;	Length 433;
Best Local Similarity	75.3%;	Pred. No. 8.4e-35;		
Matches 64;	Conservative	9;	Mismatches 12;	Indels 0;
				Gaps 0;

```

QY      1 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATV LQOTYHAHRSNALQGLGKHNYCRNP DN 60
      ||||:||||:| ||||:| |||: |||||:|||||
Db      70 KTCYGGNGSHYRGKANRDLSGRPCLAWDSPYVLKMYHAHRSDAIQGLGKHNYCRNP DN 129

```

```

Qy      61 RRRPWCYOVGKLPLVQECMNHDC 85
          :|||||||:|||| || ||| ||:
Db      130 QRRPWCYOVGKLQFVQFCMVDCC 154

```

RESULT 6

UKMS

u-plasminogen activator (EC 3.4.21.73) precursor - mouse
C/Species: Mus musculus (house mouse)
C/Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 18-Jun-1999
C/Accession: A29420; A24615

R: Degen, S.J.F.; Heckel, J.L.; Reich, E.; Degen, J.L.
Biochemistry 26, 8270-8279, 1987
A: Title: The murine urokinase-type plasminogen activator gene.
A: Reference number: A29420; MUID:88163489; PMID:2831940
A: Accession: A29420
A: Molecule type: DNA
A: Residues: 1-433 <DEG>
A: Cross-references: GB:M17922; NID:g2022296; PIDN:AAA40539.1; PID:g2022297
R: Belin, D.; Vassalli, J.D.; Combe, C.; Godeau, F.; Nagamine, Y.; Reich, E.; Kocher, Eur. J. Biochem. 148, 225-232, 1985
A: Title: Cloning, nucleotide sequencing and expression of cDNAs encoding mouse urokinase
A: Reference number: A24615; MUID:85179474; PMID:2985383
A: Accession: A24615
A: Molecule type: mRNA
A: Residues: 1-433 <BEL>
A: Cross-references: GB:X02389; NID:g55127; PIDN:CAA26231.1; PID:g55128
C: Genetics:
A: Introns: 19/3; 30/1; 66/1; 124/2; 155/1; 229/2; 279/1; 326/1; 375/3
C: Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C: Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F: 1-20/Domain: signal sequence #status predicted <SIG>
F: 21-178/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F: 32-63/Domain: EGF homology <EGF>
F: 71-152/Domain: kringle homology <KR>
F: 180-433/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
F: 180-421/Domain: trypsin homology <TRY>
F: 169-301, 211-227, 219-290, 315-384, 347-363, 374-402/Disulfide bonds: #status predicted
F: 226, 277, 378/Active site: His, Asp, Ser #status predicted

Query Match 76.4%; Score 381; DB 1; Length 433;
Best Local Similarity 72.9%; Pred. No. 1.4e-34;
Matches 62; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLOQTYHAHRSNALQLGKHNCRPN 60
Db 69 KTCYHNGDSYRGKANTDTKGRPCLANAPAVLQKRYNAHRDAISLGKHNCRPN 128
QY 61 RRPWCYVQVGLKPLVQECMVHDC 85
Db 129 QKRPWCYVQIGLRFVQECMVHDCS 153

RESULT 7
JS0599
t-plasminogen activator (EC 3.4.21.68) beta precursor - common vampire bat
N: Alternate names: tissue plasminogen activator
C: Species: Desmodus rotundus (common vampire bat)
C: Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C: Accession: JS0599
R: Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; Don
Gene 105, 229-237, 1991
A: Title: The plasminogen activator family from the salivary gland of the vampire bat Des
A: Reference number: JS0597; MUID:92039036; PMID:1937019
A: Accession: JS0599
A: Molecule type: mRNA
A: Residues: 1-431 <KRA>
A: Cross-references: GB:M63989; NID:g166076; PIDN:AAA31594.1; PID:g166077
C: Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C: Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F: 1-21/Domain: signal sequence #status predicted <SIG>
F: 22-36/Domain: propeptide #status predicted <PRO>
F: 37-431/Product: plasminogen activator beta #status predicted <PLA>
F: 41-74/Domain: EGF homology <EGF>
F: 82-163/Domain: kringle homology <KR>
F: 180-425/Domain: trypsin homology <TRY>
F: 41-52, 46-63, 65-74, 82-163, 103-145, 134-158, 168-299, 211-227, 219-288, 313-388/Disulfide bon
F: 139, 352/Binding site: carbohydrate (Asn) (covalent) #status predicted
F: 179-180/Cleavage site: His-Ser (plasmin) #status predicted
F: 226, 275, 382/Active site: His, Asp, Ser #status predicted
F: 345-361, 378-406/Disulfide bonds: #status predicted

Query Match 47.3%; Score 236; DB 2; Length 431;
Best Local Similarity 48.8%; Pred. No. 1.7e-18;

Matches 41; Conservative 13; Mismatches 30; Indels 0; Gaps 0;

QY 2 TCEYEGNGHFYRGKASTDTMGRCPLPWN SATVLOQTYHAHRSNALQLGKHNCRPN 61
Db 81 TCEYKDGVTYRGTSSESQAQICINWNSNLLTRTYNGRRSDAITLGLGNHNYCRPN 140
QY 62 RRPWCYVQVGLKPLVQECMVHDC 85
Db 141 SKPWCYVIKASKFILEFCSPVCS 164

RESULT 8
A34369
t-plasminogen activator (EC 3.4.21.68) precursor - false vampire bat (Megaderma lyra)
C: Species: Megaderma lyra
C: Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C: Accession: A34369
R: Gardell, S.J.; Duong, L.T.; Diehl, R.E.; York, J.D.; Hare, T.R.; Register, R.B.; Ja
J. Biol. Chem. 264, 17947-17952, 1989
A: Title: Isolation, characterization, and cDNA cloning of a vampire bat salivary plas
A: Reference number: A34369; MUID:90036867; PMID:2509450
A: Accession: A34369
A: Status: preliminary
A: Molecule type: mRNA
A: Residues: 1-477 <GAR>
A: Cross-references: GB:J05082; NID:g166080; PIDN:AAA31596.1; PID:g166081
C: Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat
C: Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F: 1-21/Domain: signal sequence #status predicted <SIG>
F: 22-36/Domain: propeptide #status predicted <PRO>
F: 37-477/Product: plasminogen activator #status predicted <PLA>
F: 42-79/Domain: fibronectin type I repeat homology <IFA>
F: 87-120/Domain: EGF homology <EGF>
F: 128-209/Domain: kringle homology <KR>
F: 226-471/Domain: trypsin homology <TRY>
F: 42-72, 70-79, 87-98, 92-109, 111-120, 128-209, 149-191, 180-204, 214-345, 257-273, 265-334, 35
F: 272, 321, 428/Active site: His, Asp, Ser #status predicted

Query Match 47.3%; Score 236; DB 1; Length 477;
Best Local Similarity 48.8%; Pred. No. 1.9e-18;
Matches 41; Conservative 13; Mismatches 30; Indels 0; Gaps 0;

QY 2 TCEYEGNGHFYRGKASTDTMGRCPLPWN SATVLOQTYHAHRSNALQLGKHNCRPN 61
Db 127 TCEYKDGVTYRGTSSESQAQICINWNSNLLTRTYNGRRSDAITLGLGNHNYCRPN 186
QY 62 RRPWCYVQVGLKPLVQECMVHDC 85
Db 187 SKPWCYVIKASKFILEFCSPVCS 210

RESULT 9
JS0598
t-plasminogen activator (EC 3.4.21.68) alpha-2 precursor - common vampire bat
N: Alternate names: tissue plasminogen activator
C: Species: Desmodus rotundus (common vampire bat)
C: Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C: Accession: JS0598
R: Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.;
Gene 105, 229-237, 1991
A: Title: The plasminogen activator family from the salivary gland of the vampire bat
A: Reference number: JS0597; MUID:92039036; PMID:1937019
A: Accession: JS0598
A: Molecule type: mRNA
A: Residues: 1-477 <KRA>
A: Cross-references: GB:M63988; NID:g166074; PIDN:AAA31593.1; PID:g166075
C: Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat
C: Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F: 1-21/Domain: signal sequence #status predicted <SIG>
F: 22-36/Domain: propeptide #status predicted <PRO>
F: 37-477/Product: plasminogen activator alpha-2 #status predicted <PLA>
F: 42-79/Domain: fibronectin type I repeat homology <IFA>
F: 87-120/Domain: EGF homology <EGF>

A:Experimental source: embryonic lung fibroblast IMR-90 cells
A:Note: part of this sequence, including the amino end of the mature protein, was confirmed by peptide sequencing
R:Pennington, D.; Holmes, W.E.; Kohr, W.J.; Harkins, R.N.; Vohar, G.A.; Ward, C.A.; Bennett, Nature 301, 214-221, 1983
A:Title: Cloning and expression of human tissue-type plasminogen activator cDNA in Escherichia coli
A:Reference number: A93293; MUID:83115262; PMID:6337343
A:Accession: A93293
A:Molecule type: mRNA
A:Residues: 1-562 <PEN>
A:Cross-references: GB:L00141
A:Experimental source: melanoma cells
R:Sasaki, H.; Saito, Y.; Hayashi, M.; Otsuka, K.; Niwa, M.
Nucleic Acids Res. 16, 5695, 1988
A:Title: Nucleotide sequence of the tissue-type plasminogen activator cDNA from human fetal lung
A:Reference number: S02125; MUID:88262579; PMID:3133640
A:Accession: S02125
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-562 <SAS>
A:Cross-references: EMBL:X07393; NID:g37243; PIDN:CAA30302.1; PID:g37244
A:Experimental source: fetal lung cells
R:Kagltani, H.; Tagawa, M.; Hatanaka, K.; Ikari, T.; Saito, A.; Bando, H.; Okada, K.; Maehara, T.; Miyazawa, T.; Ranby, L.O.; Palm, G.; Holmgren, E.; Josephson, S.
Proc. Natl. Acad. Sci. U.S.A. 80, 349-352, 1983
A:Title: Isolation of cDNA sequences coding for a part of human tissue plasminogen activator
A:Reference number: A93951; MUID:83169656; PMID:6572897
A:Accession: A93951
A:Molecule type: mRNA
A:Residues: 251-358 <EDL>
A:Experimental source: melanoma cells
R:Polh, G.; Kalistrom, M.; Bergsdorf, N.; Wallen, P.; Jornvall, H.
Biochemistry 23, 3701-3707, 1984
A:Title: Tissue plasminogen activator: peptide analyses confirm an indirectly derived amino acid sequence
A:Reference number: A90488; MUID:85000468; PMID:6433976
A:Contents: annotation; melanoma cells, partial sequence of residues 36-562, active and inactive
FEBS Lett. 168, 29-32, 1984
A:Title: Differences between uterine and melanoma forms of tissue plasminogen activator
A:Reference number: A91322; MUID:84158956; PMID:6538514
A:Accession: A91322
A:Molecule type: protein
A:Residues: 33-45; 311-320 <POH>
A:Experimental source: uterus
A:Note: In the uterus, cleavage of the activation peptide may also occur after 38-Gln
R:van Zonneveld, A.J.; Veerman, H.; Pannekoek, H.
J. Biol. Chem. 261, 14214-14218, 1986
A:Reference number: A37567; MUID:87033611; PMID:3021732
A:Contents: annotation; fibrin binding site
R:Verheijen, J.H.; Caspers, M.P.M.; Chang, G.T.G.; de Munk, G.A.W.; Pouwels, P.H.; Engerink, J. 5, 3525-3530, 1986
EMBO J. 5, 3525-3530, 1986
A:Title: Involvement of finger domain and kringle 2 domain of tissue-type plasminogen activator in fibrin binding
A:Reference number: A37568; MUID:87161761; PMID:3030730
A:Contents: annotation; fibrin binding site
R:Dodd, I.; Nunn, B.; Robinson, J.H.
Thromb. Haemost. 59, 523-528, 1988
A:Title: Isolation, identification and pharmacokinetic properties of human tissue-type plasminogen activator
A:Reference number: A60902; MUID:89044681; PMID:3142086
A:Contents: annotation; novel forms of expressed recombinant t-PA
R:Harris, T.J.R.; Patel, T.; Marston, F.A.O.; Little, S.; Emtage, J.S.; Opdenakker, G.; Mol. Biol. Med. 3, 279-292, 1986
A:Title: Cloning of cDNA coding for human tissue-type plasminogen activator and its expression in Escherichia coli
A:Reference number: A54645; MUID:86284200; PMID:3090401
A:Accession: A54645
A:Molecule type: mRNA
A:Residues: 1-562 <HAR>

A:Cross-references: GB:M15518; NID:g190031; PIDN:AAA6011.1; PID:g190032
A:Note: parts of this sequence were confirmed by peptide sequencing
R:Reddy, V.B.; Garramone, A.J.; Sasaki, H.; Wei, C.
DNA 6, 461-472, 1987
A:Title: Expression of human uterine tissue-type plasminogen activator in mouse cells
A:Reference number: I60110; MUID:88054470; PMID:2824147
A:Accession: I60110
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-562 <RES>
A:Cross-references: GB:M18182; NID:g340176; PIDN:AAA36800.1; PID:g340177
R:Fisher, R.; Waller, E.K.; Grossi, G.; Thompson, D.; Tizard, R.; Schleuning, W.D.
J. Biol. Chem. 260, 11223-11230, 1985
A:Title: Isolation and characterization of the human tissue-type plasminogen activator cDNA
A:Reference number: I55232; MUID:85289338; PMID:3161893
A:Accession: I55232
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-36 <RE2>
A:Cross-references: GB:M11890; NID:g339837; PIDN:AAA61213.1; PID:g339839
C:Comment: Cleavage by plasmin or trypsin produces two chains held together by a single disulfide bond
C:Comment: t-PA converts plasminogen to plasmin by hydrolyzing a single Arg-Val bond
C:Comment: t-PA binds chain A of fibrin by kringle 2 and the fibronectin type I repeat
C:Gene: GDB:PLAT
A:Cross-references: GDB:119496; OMIM:173370
A:Map position: 8p12-8p12
A:Introns: 24/3; 39/1; 85/1; 122/1; 180/2; 211/1; 268/2; 297/1; 362/2; 408/1; 455/3;
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-32/Domain: propeptide #status predicted <PRO>
F:33-562/Product: t-plasminogen activator #status experimental <MAT>
F:33-310/Product: t-plasminogen activator chain A #status experimental <ACH>
F:41-78/Domain: fibronectin type I repeat homology <1F1>
F:86-119/Domain: EGF homology <EGF>
F:127-208/Domain: kringle homology <KR1>
F:215-296/Domain: kringle homology <KR2>
F:311-562/Product: t-plasminogen activator chain B #status experimental <BCH>
F:311-556/Domain: trypsin homology <TRY>
F:41-71, 69-78, 86-97, 91-108, 110-119, 127-208, 148-190, 179-203, 215-296, 236-278, 267-291, 29
F:152, 483/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:219/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental
F:310-311/Cleavage site: Arg-Ile (plasmin, trypsin) #status experimental
F:357, 406/Active site: His, Asp #status predicted
F:513/Active site: Ser #status experimental

Query Match 43.3% Score 216; DB 1; Length 562;
Best Local Similarity 47.0%; Pred. No. 3.7e-16;
Matches 39; Conservative 9; Mismatches 35; Indels 0; Gaps 0;

QY 3 CYEGNGHYRGAQSTDTMGRCPLPWNATVLOQTYHAHNSNALQGLGKHNCRNPDNR 62
DB 215 CYFGNGSAYRGTHTSLTESGASCLPWNMSILGKVTYTAQNPSAQLGLGKHNCRNPDGDA 274

QY 63 RPYGVYQVGLKPLVQECMVHDC 85
DB 275 KPWCHVTLKNRRLTWEYCDVPSCS 297

RESULT 13
I38098
t-Plasminogen activator precursor, inactive endothelial splice form - human
N:Alternate names: tissue plasminogen activator
C:Species: Homo sapiens (man)
C>Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 22-Jun-1999
C:Accession: I38098; S01678
R:Siebert, P.D.; Fong, K.
Nucleic Acids Res. 18, 1086, 1990
A:Title: Variant tissue type plasminogen activator (PLAT) cDNA obtained from human en
A:Reference number: I38098; MUID:90192128; PMID:1969145
A:Accession: I38098
A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA
A:Residues: 1-291 <SIE>
A:Cross-references: EMBL:X13097; NID:g35282; PIDN:CAA31489.1; PID:g35283
C:Comment: For the main splice form, see PIR:UKHUT. This form probably does not have pro
C:Genetics:
A:Gene: GDB:PLAT
A:Cross-references: GDB:119496; OMIM:173370
A:Map position: 8p12-8p12
A:Introns: 24/3; 39/1; 85/1; 122/1; 180/2; 211/1; 268/2
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C:Keywords: alternative splicing; fibrinolysis; glycoprotein; kringle
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-32/Domain: propeptide #status predicted <PRO>
F:33-291/Product: t-plasminogen activator, inactive endothelial splice form #status pred
F:41-78/Domain: fibronectin type I repeat homology <1FA>
F:86-119/Domain: EGF homology <EGF>
F:127-208/Domain: kringle homology <KR1>
F:215-291/Domain: kringle homology #status atypical <KR2>
F:41-71,69-78,86-97,91-108,110-119,127-208,148-190,179-203/Disulfide bonds: #status pred

Query Match 43.1%; Score 215; DB 2; Length 291;
Best Local Similarity 45.9%; Pred. No. 2,5e-16;
Matches 39; Conservative 9; Mismatches 37; Indels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQQT YHAHRSNALQLGKHNCRNPDR 61
DB 126 TCYEDQGISYRGRTWSTAESGAECTWNSSALQANAYSGRRPD ARLGLGNHNYCRNPDRD 185
QY 62 RRPWCYVQVGLKPLVQECMVHDCAD 86
DB 186 SKPWCYVFKA GKYSSSEFCSTPACSE 210

RESULT 14
A29941
t-plasminogen activator (EC 3.4.21.68) precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A29941; S48205; S48207; S48206
R:Rickles, R.J.; Darow, A.L.; Strickland, S.
J. Biol. Chem. 263, 1563-1569, 1988
A:Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator m
A:Reference number: A29941; MUID:88087303; PMID:2826484
A:Accession: A29941
A:Molecule type: mRNA
A:Residues: 1-559 <RTC>
A:Cross-references: GB:J03520; NID:g202109; PIDN:AAA40470.1; PID:g202110
R:Lijnen, H.R.; van Hoef, B.; Beelen, V.; Collen, D.
Eur. J. Biochem. 224, 863-871, 1994
A:Title: Characterization of the murine plasma fibrinolytic system.
A:Reference number: S48202; MUID:95010076; PMID:7523120
A:Accession: S48205
A:Molecule type: protein
A:Residues: 33-37,'X',39-40 <LIJ>
A:Accession: S48207
A:Molecule type: protein
A:Residues: 309-316 <LI2>
A:Accession: S48206
A:Molecule type: protein
A:Residues: 33-37,'X',39-40 <LIW>
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-29/Domain: propeptide #status predicted <PRO>
F:30-559/Product: t-plasminogen activator #status predicted <MAT>
F:30-308/Product: t-plasminogen activator chain A #status predicted <ACH>
F:38-75/Domain: fibronectin type I repeat homology <1FI>
F:83-116/Domain: EGF homology <EGF>
F:124-205/Domain: kringle homology <KR1>
F:213-294/Domain: kringle homology <KR2>
F:309-559/Product: t-plasminogen activator chain B #status predicted <BCH>
F:309-553/Domain: trypsin homology <TRY>
F:38-68,66-75,83-94,88-105,107-116,124-205,145-187,176-200,213-294,234-276,265-289,297-4

F:149,481/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:308-309/Cleavage site: Arg-Ile (plasmin, trypsin) #status predicted
F:355,404,510/Active site: His, Asp, Ser #status predicted

Query Match 43.1%; Score 215; DB 1; Length 559;
Best Local Similarity 47.0%; Pred. No. 4,7e-16;
Matches 39; Conservative 10; Mismatches 34; Indels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQQT YHAHRSNALQLGKHNCRNPDR 61
DB 123 TCYEDQGITRGTWSTAESGAECIMNNSVLSKPYNARRPNAIKLGLGNHNYCRNPDRD 182
QY 62 RRPWCYVQVGLKPLVQECMVHDC 84
DB 183 LRPWCYVFKA GKYTTEFCSTPAC 205

RESULT 15
JS0597
t-plasminogen activator (EC 3.4.21.68) alpha-1 precursor - common vampire bat
N:Alternate names: tissue plasminogen activator
C:Species: Desmodus rotundus (common vampire bat)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C:Accession: JS0597
R:Kraetzschmar, J.; Haendler, B.; Langer, G.; Boldol, W.; Bringmann, P.; Alagon, A.;
Gene 105, 229-237, 1991
A:Title: The plasminogen activator family from the salivary gland of the vampire bat
A:Reference number: JS0597; MUID:92039036; PMID:1937019
A:Accession: JS0597
A:Molecule type: mRNA
A:Residues: 1-477 <KRA>
A:Cross-references: GB:M63987; NID:g166070; PIDN:AAA31591.1; PID:g166071
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-36/Domain: propeptide #status predicted <PRO>
F:37-477/Product: plasminogen activator alpha-1 #status predicted <PLA>
F:42-79/Domain: fibronectin type I repeat homology <1FA>
F:87-120/Domain: EGF homology <EGF>
F:128-209/Domain: kringle homology <KRG>
F:226-471/Domain: trypsin homology <TRY>
F:42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,35
F:153,398/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:225-226/Cleavage site: His-Ser (plasmin) #status predicted
F:272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 41.7%; Score 208; DB 2; Length 477;
Best Local Similarity 45.2%; Pred. No. 2,4e-15;
Matches 38; Conservative 11; Mismatches 35; Indels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQQT YHAHRSNALQLGKHNCRNPDR 61
DB 127 TCYEGQVTYRGTWSTAESRYECIMNNSLTRYNGRMPDAFNLGLGNHNYCRNPNGA 186
QY 62 RRPWCYVQVGLKPLVQECMVHDC 85
DB 187 KRPWCYVIKAGKFTSESCSVPVCS 210

Search completed: November 11, 2002, 13:02:46
Job time : 39 secs

100

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2002, 11:40:46 ; Search time 21 Seconds

(without alignments)
169.855 Million cell updates/sec

Title: PCT-US02-27855-1

Perfect score: 499

Sequence: 1 KTCYEGNGHFYRGKASTDTM.....YVQVGLKPLVQECMVHDCAD 86

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	494	99.0	431	1 UROK_HUMAN	P00749 homo sapien
2	459	92.0	433	1 UROK_PAPCY	P16227 papio cynoc
3	409	82.0	442	1 UROK_PIG	P04185 sus scrofa
4	400	80.2	432	1 UROK_RAT	P29598 ratus norv
5	383	76.8	433	1 UROK_BOVIN	Q05589 bos taurus
6	381	76.4	433	1 UROK_MOUSE	P06869 mus musculu
7	236	47.3	431	1 URTB_DESRO	P98121 desmodus ro
8	236	47.3	477	1 URT2_DESRO	P15638 desmodus ro
9	226	45.3	559	1 TPA_RAT	P19637 ratus norv
10	218	43.7	394	1 URTG_DESRO	P19150 desmodus ro
11	216	43.3	562	1 TPA_HUMAN	P00750 homo sapien
12	215	43.1	559	1 TPA_MOUSE	P12124 mus musculu
13	214	42.9	566	1 TPA_BOVIN	Q28198 bos taurus
14	208	41.7	477	1 URT1_DESRO	P98119 desmodus ro
15	203.5	40.8	434	1 UROK_CHICK	P15120 gallus gall
16	193	38.7	655	1 HGFA_HUMAN	Q04756 homo sapien
17	188	37.7	653	1 HGFA_MOUSE	Q91098 mus musculu
18	179.5	36.0	603	1 FA12_CAVPO	Q04962 cavia porce
19	170.5	34.2	615	1 FA12_HUMAN	P00748 homo sapien
20	158	31.7	473	1 KREM_MOUSE	Q09n43 mus musculu
21	158	31.7	473	1 KREM_RAT	Q92454 ratus norv
22	156	31.3	1420	1 APOA_MACMU	P14417 macaca mula
23	155.5	31.2	593	1 FA12_BOVIN	P98140 bos taurus
24	155	31.1	475	1 KREM_HUMAN	Q06m08 homo sapien
25	155	31.1	4548	1 APOA_HUMAN	P08519 homo sapien
26	154	30.9	812	1 PLMN_BOVIN	P06868 bos taurus
27	153	30.7	790	1 PLMN_PIG	P06867 sus scrofa
28	152.5	30.6	169	1 PLMN_RAT	Q01177 ratus norv
29	149	29.9	810	1 PLMN_HUMAN	P00747 homo sapien
30	148.5	29.8	944	1 ROR2_MOUSE	Q92138 mus musculu
31	147.5	29.6	810	1 PLMN_ERIEU	Q29485 erinaceus e
32	147	29.5	937	1 ROR1_HUMAN	Q01973 homo sapien
33	147	29.5	937	1 ROR1_MOUSE	Q92139 mus musculu

34	146	29.3	810	1 PLMN_MACMU	P12545 macaca mula
35	143.5	28.8	711	1 HGFL_HUMAN	P26927 homo sapien
36	143.5	28.8	943	1 ROR2_HUMAN	Q01974 homo sapien
37	143	28.7	333	1 PLMN_CANFA	P80009 canis fam1
38	137	27.5	812	1 PLMN_MOUSE	P20918 mus musculu
39	133.5	26.8	728	1 HGF_MOUSE	Q08048 mus musculu
40	131.5	26.4	716	1 HGFL_MOUSE	P26928 mus musculu
41	131.5	26.4	728	1 HGF_HUMAN	P14210 homo sapien
42	131.5	26.4	728	1 HGF_RAT	P17945 ratus norv
43	126	25.3	325	1 PLMN_PETMA	P33574 petromyzon
44	125	25.1	622	1 THRB_HUMAN	P00734 homo sapien
45	124.5	24.9	618	1 THRB_MOUSE	P19221 mus musculu

ALIGNMENTS

RESULT 1

UROK_HUMAN STANDARD; PRT; 431 AA.

ID UROK_HUMAN

AC P00749; Q15844; Q16618; Q969W6;

DT 21-JUL-1986 (Rel. 01, Created)

DT 20-MAR-1987 (Rel. 04, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (UPA)

(U-plasminogen activator).

GN PLAU.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=85215647; PubMed=2987867;

RA Riccio A., Grimaldi G., Verde P., Sebastio G., Boast S., Blas F.;

RT "The human urokinase-plasminogen activator gene and its promoter.";

RL Nucleic Acids Res. 13:2759-2771(1985).

RN [2]

RP SEQUENCE FROM N.A.

RA Holmes W.E., Pennica D., Blaber M., Rey M.W., Guenzler W.A.,

RT Steffens G.J., Heyneker H.L.;

RT "Cloning and expression of the gene for pro-urokinase in Escherichia

RL coli.";

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE=86056954; PubMed=2415429;

RA Nagai M., Hiratsutsu R., Kaneda T., Hayasuke N., Arimura H.,

RT Nishida M., Suyama T.;

RT "Molecular cloning of cDNA coding for human prepro-urokinase.";

RL Gene 36:183-188(1985).

RN [4]

RP SEQUENCE FROM N.A.

RX MEDLINE=85203359; PubMed=3888571;

RA Jacobs P., Crayador A., Loriau R., Brockly F., Colau B., Chuchana P.,

RT van Elsen A., Herzog A., Bollen A.;

RT "Molecular cloning, sequencing, and expression in Escherichia coli of

RL human prepro-urokinase cDNA.";

RL DNA 4:139-146(1985).

RN [5]

RP SEQUENCE FROM N.A.

RA Riederer M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,

RT Nickerson D.A.;

RT Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.

RN [6]

RP SEQUENCE FROM N.A.

RC TISSUE=Lung;

RA Strausberg R.;

RT Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

RN [7]

RP SEQUENCE OF 66-431 FROM N.A.

RX MEDLINE=84272706; PubMed=6589620;

RL Verde P., Stoppelli M.P., Galeffi P., di Nocera P., Blas F.;

RT "Identification and primary sequence of an unspliced human urokinase
RP poly(A)+ RNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:4727-4731(1984).
RN [8]
RP SEQUENCE OF 21-177.
RX MEDLINE=83055084; Pubmed=6754569;
RA Gunzler W.A., Steffens G.J., Otting F., Kim S.-M.A., Frankus E.,
RN Flohe L.;
RT "The primary structure of high molecular mass urokinase from human
RL urine. The complete amino acid sequence of the A chain.";
RN Hoppe-Seyler's Z. Physiol. Chem. 363:1155-1165(1982).
RP SEQUENCE OF 156-176 AND 179-224.
RX MEDLINE=83003608; Pubmed=6749491;
RA Schaller J., Nick H., Rickli E.E., Gillesen D., Lergier W.,
RN Studer R.O.;
RT "Human low-molecular-weight urinary urokinase. Partial
RT characterization and preliminary sequence data of the two polypeptide
RL chains.";
RN Eur. J. Biochem. 125:251-257(1982).
RP SEQUENCE OF 158-410.
RX MEDLINE=83055099; Pubmed=6754572;
RA Steffens G.J., Gunzler W.A., Otting F., Frankus E., Flohe L.;
RT "The complete amino acid sequence of low molecular mass urokinase
RT from human urine.";
RN Hoppe-Seyler's Z. Physiol. Chem. 363:1043-1058(1982).
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX MEDLINE=96000858; Pubmed=8591045;
RA Spraggon G., Phillips C., Nowak U.K., Ponting C.P., Saunders D.,
RN Dobson C.M., Stuart D.I., Jones E.Y.;
RT "The crystal structure of the catalytic domain of human
RT urokinase-type plasminogen activator.";
RL Structure 3:681-691(1995).
RN [12]
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 159-411.
RX MEDLINE=20266327; Pubmed=10805774;
RA Sperl S., Jacob U., Arroyo de Prada N., Sturzebecher J., Wilhelm O.G.,
RN Bode W., Magdolen V., Huber R., Moroder L.;
RT "(4-aminomethyl)phenylguanidine derivatives as nonpeptidic highly
RT selective inhibitors of human urokinase.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:5113-5118(2000).
RN [13]
RP STRUCTURE BY NMR.
RX MEDLINE=89127526; Pubmed=2536903;
RA Oswald R.E., Bogusky M.J., Bamberger M., Smith R.A.G., Dobson C.M.;
RT "Dynamics of the multidomain fibrinolytic protein urokinase from two-
RT dimensional NMR.";
RL Nature 337:579-582(1989).
RN [14]
RP STRUCTURE BY NMR OF 67-155.
RX MEDLINE=93003110; Pubmed=1327118;
RA Li X., Smith R.A.G., Dobson C.M.;
RT "Sequential 1H NMR assignments and secondary structure of the kringle
RT domain from urokinase.";
RN Biochemistry 31:9562-9571(1992).
RP STRUCTURE BY NMR OF 67-155.
RX MEDLINE=94149701; Pubmed=8107091;
RA Li X., Bokman A.M., Llinas M., Smith R.A.G., Dobson C.M.;
RT "Solution structure of the kringle domain from urokinase-type
RT plasminogen activator.";
RN J. Mol. Biol. 235:1548-1559(1994).
RP VARIANT LEU-141.
RX MEDLINE=96186279; Pubmed=8652631;
RA Yoshimoto M., Ushiyama Y., Sakai M., Tamaki S., Hara H., Takahashi K.,
RN Sawasaki Y., Hanada K.;
RT "Characterization of single chain urokinase-type plasminogen
RT activator with a novel amino-acid substitution in the kringle
RL structure.";
RN Biochim. Biophys. Acta 1293:83-89(1996).

RN [17]
RP VARIANT LEU-141.
RX MEDLINE=97218551; Pubmed=9065988;
RA Conne B., Berczy M., Belin D.;
RT "Detection of polymorphisms in the human urokinase-type plasminogen
RT activator gene.";
RN Thromb. Haemost. 77:434-435(1997).
RP ERRATUM.
RX Conne B., Berczy M., Belin D.;
RN Thromb. Haemost. 78:973-973(1997).
RP VARIANT LEU-141.
RX MEDLINE=97337920; Pubmed=9194591;
RA Turkmen B., Schmitt M., Schmalfeldt B., Trommler P., Hell W.,
RN Creutzburg S., Graeff H., Magdolen V.;
RT "Mutational analysis of the genes encoding urokinase-type plasminogen
RT activator (uPA) and its inhibitor PAI-1 in advanced ovarian cancer.";
RL Electrophoresis 18:686-689(1997).
CC -1- FUNCTION: POTENT PLASMINOGEN ACTIVATOR AND IS CLINICALLY USED FOR
CC THERAPY OF THROMBOLYTIC DISORDERS.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW
CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN.
CC -1- PHARMACEUTICAL: Available under the name Abbokinase (Abbott). Used
CC in pulmonary embolism (PE) to initiate fibrinolysis.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X02419; CAA26268.1; -
DR EMBL; M15476; AAA61253.1; -
DR EMBL; D00244; BAA00175.1; -
DR EMBL; D11143; BAA01919.1; -
DR EMBL; X02760; CAA26535.1; -
DR EMBL; AF377330; AAK53822.1; -
DR EMBL; BC013575; AAH13575.1; -
DR EMBL; K03226; AAC97138.1; -
DR EMBL; K02286; AAA61252.1; -
DR EMBL; A21571; CAA01559.1; -
DR EMBL; A18397; CAA01390.1; -
DR PIR; A00931; UKHU.
DR PIR; A32974; A32974.
DR PDB; 1KDU; 31-OCT-93.
DR PDB; 1LMW; 29-JAN-96.
DR PDB; 1URK; 08-MAY-95.
DR PDB; 1EJN; 17-MAY-00.
DR MEROPS; S01.231; -
DR GlycoSuiteDB; P00749; -
DR Genew; HGNC:9052; PLA0.
DR MIM; 191840; -
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser-protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.

Query Match 99.0%; Score 494; DB 1; Length 431;
Best Local Similarity 98.8%; Pred. No. 8.1e-50;
Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQGLGKHNYCRNPDN 60
D6 68 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQGLGKHNYCRNPDN 127

QY 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
D6 128 RRRPWCYVQVGLKPLVQECMVHDCAD 153

RESULT 2
UROK_PAPCY STANDARD; PRT; 433 AA.
ID UROK_PAPCY STANDARD; PRT; 433 AA.
AC P16227;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLAU.
OS Papio cynocephalus (Yellow baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Papio.
OX NCBI_TaxID=9556;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thoracic aorta;
RX MEDLINE=90287734; PubMed=2113276;
RA Au Y.P.T., Wang T.W., Clowes A.W.;
RT "Nucleotide and deduced amino acid sequences of baboon urokinase-type
RT plasminogen activator."
RL Nucleic Acids Res. 18:3411-3411(1990).
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW
CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X51935; CAA36200.1; -.
DR PIR; S14687; UKBAY.
DR HSSP; P00749; LLMW.
DR MEROPS; S01.231; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRODOM; PD000395; kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRYPSIN; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE_NEG.
DR PROSITE; PS00021; KRINGLE_1; 1.

DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Plasminogen activation; Hydrolyase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20
FT CHAIN 21 433
FT CHAIN 21 176
FT CHAIN 155 176
FT CHAIN 178 433
FT CHAIN 26 62
FT DOMAIN 69 150
FT DOMAIN 151 177
FT DOMAIN 178 433
FT DOMAIN 30 38
FT DISULFID 32 50
FT DISULFID 52 61
FT DISULFID 167 298
FT DISULFID 208 224
FT DISULFID 216 287
FT DISULFID 315 384
FT DISULFID 347 363
FT DISULFID 374 402
FT ACT_SITE 223 223
FT ACT_SITE 274 274
FT ACT_SITE 378 378
FT CARBOHYD 324 324
SQ SEQUENCE 433 AA; 48595 MW; 816D22DFEDDC8792 CRG64;
Query Match 92.0%; Score 459; DB 1; Length 433;
Best Local Similarity 93.0%; Pred. No. 9.2e-46;
Matches 80; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQGLGKHNYCRNPDN 60
D6 67 KTCYEGNGHFYRGKASTDTMGRSCLAWNSATVLOQTYHAHRSNALQGLGKHNYCRNPDN 126

QY 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
D6 127 RRRPWCYVQVGLKPLVQECMVHDCAD 152

RESULT 3
UROK_PIG STANDARD; PRT; 442 AA.
ID UROK_PIG STANDARD; PRT; 442 AA.
AC P04185;
DT 20-MAR-1987 (Rel. 04, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLAU.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=85087954; PubMed=6096832;
RA Nagamine Y., Pearson D., Altus M.S., Reich E.;
RT "cDNA and gene nucleotide sequence of porcine plasminogen activator."
RL Nucleic Acids Res. 12:9525-9541(1984).
RN [2]
RP REVISION TO 241.
RA Nagamine Y.;
RL Submitted (DEC-1986) to the PIR data bank.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; X01648; CAA25806.1; -
DR EMBL; X02724; CAA26511.1; -
DR PIR; A00932; UKPG.
DR HSSP; P00749; IKDU.
DR MEROPS; S01.231; -
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser-protease_Try.
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE_NEG.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Plasminogen activation; Hydrolyase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; zymogen; signal.
FT SIGNAL 1 20
FT CHAIN 21 442
FT CHAIN 21 188
FT CHAIN 190 442
FT DOMAIN 29 65
FT DOMAIN 72 153
FT DOMAIN 154 189
FT DOMAIN 190 442
FT CARBOHYD 152 152
FT DISULFID 33 41
FT DISULFID 35 53
FT DISULFID 55 64
FT DISULFID 179 310
FT DISULFID 220 236
FT DISULFID 228 299
FT DISULFID 324 393
FT DISULFID 356 372
FT DISULFID 383 411
FT ACT_SITE 235 235
FT ACT_SITE 286 286
FT ACT_SITE 387 387
FT CONFLICT 241 241
FT CONFLICT 242 242
FT CONFLICT 288 288
SQ SEQUENCE 442 AA; 49116 MW; EE32FCEFS0132IEE CRC64;

Query Match 82.0%; Score 409; DB 1; Length 442;
Best Local Similarity 82.4%; Pred. No. 5.8e-40;
Matches 70; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

OY 1 KTCYEGNHGFYRGKASTDTMGRPLPWSATVLTQTYHAHRSNALQLGKHNCRNDN 60
DB 70 QTCFEGNGHSHYRGKANTNTGRCPLPWSATVLTNTYHAHRDALQLGKHNCRNDN 129
OY 61 RRRPWCYVQVGLKPLVQECMVHDC 85
DB 130 QRRPWCYVQVGLKPLVQECMVHDC 154

RESULT 4

UROK_RAT
ID UROK_RAT STANDARD; PRT: 432 AA.
AC P29598;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (UPA)
DE (U-plasminogen activator).
GN PLAU.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Fischer 344;
RX MEDLINE=92233409; PubMed=1568219;
RA Henderson B.R., Tansey W.P., Phillips S.M., Ramshaw I.A.,
RA Keiford R.F.;
RT "Transcriptional and posttranscriptional activation of urokinase
RT plasminogen activator gene expression in metastatic tumor cells.";
RT Cancer Res. 52:2489-2496(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Rabbani S.A.;
RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-1-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH
CC CONSISTS OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM
CC CONTAINS A LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 156 IN THE
CC LOW MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X63434; CAA45028.1; -
DR EMBL; X65651; CAA46601.1; -
DR PIR; S18932; S18932.
DR HSSP; P00749; IKDU.
DR MEROPS; S01.231; -
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser-protease_Try.
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE_NEG.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolyase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; zymogen; signal.
FT SIGNAL 1 19
FT CHAIN 20 432
FT CHAIN 20 177
UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
CHAIN A (BY SIMILARITY).

```
FT CHAIN 156 177 SHORT A CHAIN (A1) (BY SIMILARITY).
FT CHAIN 179 432 CHAIN B (BY SIMILARITY).
FT DOMAIN 27 63 EGF-LIKE.
FT DOMAIN 70 151 KRINGLE.
FT DOMAIN 152 178 CONNECTING PEPTIDE.
FT DOMAIN 179 432 SERINE PROTEASE.
FT DISULFID 31 39 BY SIMILARITY.
FT DISULFID 33 51 BY SIMILARITY.
FT DISULFID 53 62 BY SIMILARITY.
FT DISULFID 168 300 INTERCHAIN (BY SIMILARITY).
FT DISULFID 210 226 BY SIMILARITY.
FT DISULFID 218 289 BY SIMILARITY.
FT DISULFID 314 383 BY SIMILARITY.
FT DISULFID 346 362 BY SIMILARITY.
FT DISULFID 373 401 BY SIMILARITY.
FT ACT_SITE 225 225 CHARGE RELAY SYSTEM.
FT ACT_SITE 276 276 CHARGE RELAY SYSTEM.
FT ACT_SITE 377 377 CHARGE RELAY SYSTEM.
FT ACT_SITE 16 16 N -> H (IN REF. 2).
FT CONFLICT 24 24 E -> G (IN REF. 2).
FT CONFLICT 332 332 D -> N (IN REF. 2).
SQ SEQUENCE 432 AA; 47957 MW; 4EB1B96C716244C8 CRC64;

Query Match 80.28; Score 400; DB 1; Length 432;
Best Local Similarity 80.08; Pred. No. 6.2e-39;
Matches 68; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGKHNCRNPDN 60
    ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 68 KTCYHNGQSYRGKANTDTKGRCLAWNSPAVLQGTYNHRSDALSLGKHNCRNPDN 127

QY 61 RRRPWCYVQVGLKPLVQECMVHDC A 85
    :||||| ||| ||||| |||
Db 128 QRRPWCYVQIGLKQFVQECMVQDCS 152

RESULT 5
UROK_BOVIN STANDARD; PRT; 433 AA.
ID UROK_BOVIN
AC Q05589; Q28209;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLAU.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_Taxid-9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Aortic endothelium;
RX MEDLINE-93216119; PubMed-8385052;
RA Kraetzschmar J., Haendler B., Kojima S., Rifkin D.B.,
RA Schleuning W.-D.;
RT "Bovine urokinase-type plasminogen activator and its receptor:
RT cloning and induction by retinoic acid.";
RL Gene 125:177-183(1993).
RN [2]
RP SEQUENCE OF 12-433 FROM N.A.
RC TISSUE-Kidney;
RA Ravn P., Berglund L., Petersen T.E.;
RT "Cloning and characterization of the bovine plasminogen activators uPA
RT and tPA.";
RL Int. Dairy J. 5:605-617(1995).
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-1-Val bond in
CC plasminogen to form plasmin.
CC -1- INDUCTION: BY RETINOIC ACID.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
```

```
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L03546; AAA51419.1; -.
DR EMBL; X85801; CAA59796.1; -.
DR PIR; JN0560; JN0560.
DR HSSP; P00749; 1LMW.
DR MEROPS; S01.231; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser-protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR ProDom; PD000395; kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM0020; TRYP_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE_NEG.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW kringle; EGF-like domain; Signal; Zymogen.
FT SIGNAL 1 20
FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 179 CHAIN A (BY SIMILARITY).
FT CHAIN 181 433 CHAIN B (BY SIMILARITY).
FT DOMAIN 29 65 EGF-LIKE.
FT DOMAIN 72 153 KRINGLE.
FT DOMAIN 154 180 CONNECTING PEPTIDE.
FT DOMAIN 181 433 SERINE PROTEASE.
FT DISULFID 33 41 BY SIMILARITY.
FT DISULFID 35 53 BY SIMILARITY.
FT DISULFID 55 64 BY SIMILARITY.
FT DISULFID 170 301 INTERCHAIN (BY SIMILARITY).
FT DISULFID 211 227 BY SIMILARITY.
FT DISULFID 219 290 BY SIMILARITY.
FT DISULFID 315 384 BY SIMILARITY.
FT DISULFID 347 363 BY SIMILARITY.
FT DISULFID 374 402 BY SIMILARITY.
FT ACT_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 277 277 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 378 378 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CONFLICT 189 189 A -> T (IN REF. 2).
SQ SEQUENCE 433 AA; 48730 MW; 4DE1B8D4DA47027A CRC64;

Query Match 76.8%; Score 383; DB 1; Length 433;
Best Local Similarity 75.38; Pred. No. 5.8e-37;
Matches 64; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGKHNCRNPDN 60
    ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 70 KTCYQNGHSHYRGKANDLSGRPCLAWSPTVLLKMYHAHRSDAIQLGKHNCRNPDN 129

QY 61 RRRPWCYVQVGLKPLVQECMVHDC A 85
    :||||| ||| ||||| |||
Db 130 QRRPWCYVQIGLKQFVQECMVQDCS 154

RESULT 6
UROK_MOUSE STANDARD; PRT; 433 AA.
ID UROK_MOUSE
AC P06869;
```



```

DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, last sequence update)
DT 15-JUN-2002 (Rel. 41, last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (UPA)
DE (U-plasminogen activator).
GN PLAU.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85179474; PubMed=2985383;
RA Belli D., Vassalli J.-D., Combepine C., Godeau F., Nagamine Y.,
RA Reich E., Kocher H.P., Duvolsin R.M.;
RT "Cloning, nucleotide sequencing and expression of cDNAs encoding
RT mouse urokinase-type plasminogen activator.";
RL Eur. J. Biochem. 148:225-232(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88163489; PubMed=2831940;
RA Degen S.J.F., Heckel J.L., Reich E., Degen J.L.;
RT "The murine urokinase-type plasminogen activator gene.";
RL Biochemistry 26:8270-8279(1987).
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-1-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 156 IN THE LOW
CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X02389; CAA26231.1; -.
DR EMBL: M17922; AAA40539.1; -.
DR PIR: A24615; UKMS.
DR PIR: A29420; A29420.
DR HSSP: P00749; IKDU.
DR MEROPS: S01.231; -.
DR MGD: MGI:97611; Plau.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00051; kringle; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR ProDom: PD000395; kringle; 1.
DR SMART: SM00181; EGF; 1.
DR SMART: SM00130; KR; 1.
DR SMART: SM00020; TRYP_SPC; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; FALSE_NEG.
DR PROSITE: PS00021; KRINGLE_1; 1.
DR PROSITE: PS50070; KRINGLE_2; 1.
DR PROSITE: PS50240; TRYPsin_DOM; 1.
DR PROSITE: PS00134; TRYPsin_HIS; FALSE_NEG.
DR PROSITE: PS00135; TRYPsin_SER; 1.
KW plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20
FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 178 CHAIN A (BY SIMILARITY).
FT CHAIN 157 178 SHORT A CHAIN (A1).

```

```

FT CHAIN 180 433 CHAIN B (BY SIMILARITY).
FT DOMAIN 28 64 EGF-LIKE.
FT DOMAIN 71 152 KRINGLE.
FT DOMAIN 153 179 CONNECTING PEPTIDE.
FT DOMAIN 180 433 SERINE PROTEASE.
FT DISULFID 32 40 BY SIMILARITY.
FT DISULFID 34 52 BY SIMILARITY.
FT DISULFID 54 63 BY SIMILARITY.
FT DISULFID 169 301 INTERCHAIN (BY SIMILARITY).
FT DISULFID 211 227 BY SIMILARITY.
FT DISULFID 219 290 BY SIMILARITY.
FT DISULFID 315 384 BY SIMILARITY.
FT DISULFID 347 363 BY SIMILARITY.
FT DISULFID 374 402 BY SIMILARITY.
FT ACT_SITE 226 226 CHARGE RELAY SYSTEM.
FT ACT_SITE 277 277 CHARGE RELAY SYSTEM.
FT ACT_SITE 378 378 CHARGE RELAY SYSTEM.
SQ SEQUENCE 433 AA; 48268 MW; A99C35FE250443F9 CRC64;

Query Match 76.4%; Score 381; DB 1; Length 433;
Best Local Similarity 72.9%; Pred. No. 9.9e-37;
Matches 62; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

QY 1 KTCYEKGCHFYRGKASTDTMGRPLCPWNSATVLOQTYHAHRSNALDLGLGKHNYCRNPDN 60
    ||||| ||| ||||||| ||| ||||| ||| : ||| : ||| ||||||| |||
Db 69 KTCYHGNGDSYRKGANTDTKGRPCLAWNAPAVLQKPYNHRPRDAISLGLGKHNYCRNPDN 128
    ::||| ||||| ||| : ||||| |||

QY 61 RRRPCYVOVGKLPLVQECMVHDCA 85
Db 129 QKRPCYVQIGLRQFVQECMVHDCS 153

RESULT 7
URTB_DESRO
ID URTB_DESRO STANDARD; PRT; 431 AA.
AC P98121;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator beta precursor (EC 3.4.21.68) (DSPA
beta).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OX NCBI_TaxID=9430;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=92039036; PubMed=1937019;
RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
RA Alagon A., Donner P., Schleuning W.D.;
RT "The plasminogen activator family from the salivary gland of the
RT vampire bat Desmodus rotundus: cloning and expression.";
RT Gene 105:229-237(1991).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=93393059; PubMed=1309059;
RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
RA Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common
RT vampire bat): unique fibrin specificity.";
RT Ann. N.Y. Acad. Sci. 667:395-403(1992).
CC -1- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
CC AGENT.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: MONOMER.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; M63989; AAA31594.1; -.
DR HSSP; P98119; 1A5I.
DR MEROPS; S01.239; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRYP_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolyase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Multigene family.
FT SIGNAL 1 36 POTENTIAL.
FT CHAIN 37 431 SALIVARY PLASMINOGEN ACTIVATOR BETA.
FT DOMAIN 37 75 EGF-LIKE.
FT DOMAIN 82 163 KRINGLE.
FT ACT_SITE 179 431 SERINE PROTEASE.
FT ACT_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 275 275 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 382 382 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 41 52 BY SIMILARITY.
FT DISULFID 46 63 BY SIMILARITY.
FT DISULFID 65 74 BY SIMILARITY.
FT DISULFID 82 163 BY SIMILARITY.
FT DISULFID 103 145 BY SIMILARITY.
FT DISULFID 134 158 BY SIMILARITY.
FT DISULFID 168 299 BY SIMILARITY.
FT DISULFID 211 227 BY SIMILARITY.
FT DISULFID 219 288 BY SIMILARITY.
FT DISULFID 313 388 BY SIMILARITY.
FT DISULFID 345 361 BY SIMILARITY.
FT DISULFID 378 406 BY SIMILARITY.
FT CARBOHYD 139 139 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 431 AA; 48221 MW; 699B5E675B162CBF CRC64;

Query Match 47.3%; Score 236; DB 1; Length 431;
Best Local Similarity 48.8%; Pred. No. 6.1e-20;
Matches 41; Conservative 13; Mismatches 30; Indels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASTDTMGRCLPWN SATV LQGT YHAHRSNALQLGLGKHNYCRNPDNR 61
Db 81 TCYKDGVTYRGTSSTESGACINWNSNLTTRTYNGRSDAITLGLGNHNYCRNPDNN 140
QY 62 RRPWCYVQVGLKPLVQECMVHDC A 85
Db 141 SKPMCYVIKASKFILEFCVYVCS 164

RESULT 8
URT2_DESRO

ID URT2_DESRO STANDARD; PRT; 477 AA.
AC P15638;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator alpha 2 precursor (EC 3.4.21.68) (DSPA
DE alpha-2) (BAT-PA) (T-plasminogen activator).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OX NCBI_TaxID=9430;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=92039036; PubMed=1937019;
RA Kraetzschmar J., Haendler B., Langer G., Boldol W., Bringmann P.,
RA Alagon A., Donner P., Schleuning W.D.;
RT "The plasminogen activator family from the salivary gland of the
RT vampire bat Desmodus rotundus: cloning and expression.";
RL Gene 105:229-237(1991).
RN [2]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Salivary gland;
RX MEDLINE=90036867; PubMed=2509450;
RA Gardell S.J., Duong L.T., Diehl R.E., York J.D., Hare T.R.,
RA Register R.B., Jacobs J.W., Dixon R.A.F., Friedman P.A.;
RT "Isolation, characterization, and cDNA cloning of a vampire bat
RT salivary plasminogen activator.";
RL J. Biol. Chem. 264:17947-17952(1989).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=93393059; PubMed=1309059;
RA Schleuning W.-D., Alagon A., Boldol W., Bringmann P., Petri T.,
RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Wilt W.,
RA Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common
RT vampire bat): unique fibrin specificity.";
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
CC -1- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
CC AGENT.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-I-Val bond in
CC plasminogen to form plasmin.
CC -1- ENZYME REGULATION: ACTIVITY TOWARD PLASMINOGEN IS STIMULATED IN
CC THE PRESENCE OF FIBRIN I.
CC -1- SUBUNIT: MONOMER.
CC -1- DOMAIN: THE FIBRONECTIN TYPE-I DOMAIN MEDIATES BINDING TO FIBRIN,
CC AND THE KRINGLE DOMAIN APPARENTLY MEDIATES FIBRIN-INDUCED
CC STIMULATION OF ACTIVITY.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; M63988; AAA31593.1; -.
DR EMBL; J05082; AAA31596.1; -.
DR PIR; A34369; A34369.
DR HSSP; P98119; 1A5I.
DR MEROPS; S01.232; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.

DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; Fnl; 1.
DR Pfam; PF00051; Kring1; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kring1; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; Fnl; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01253; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
Kring1; EGF-like domain; Signal; Multigene family.
FT SIGNAL 1 36 POTENTIAL.
FT CHAIN 37 477 SALIVARY PLASMINOGEN ACTIVATOR ALPHA 2.
FT DOMAIN 40 82 FIBRONECTIN TYPE-I.
FT DOMAIN 83 121 EGF-LIKE.
FT DOMAIN 128 209 KRINGLE.
FT DOMAIN 225 477 SERINE PROTEASE.
FT ACT_SITE 272 272 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 321 321 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 428 428 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 42 72 BY SIMILARITY.
FT DISULFID 70 79 BY SIMILARITY.
FT DISULFID 87 98 BY SIMILARITY.
FT DISULFID 92 109 BY SIMILARITY.
FT DISULFID 111 120 BY SIMILARITY.
FT DISULFID 128 209 BY SIMILARITY.
FT DISULFID 149 191 BY SIMILARITY.
FT DISULFID 180 204 BY SIMILARITY.
FT DISULFID 214 345 BY SIMILARITY.
FT DISULFID 257 273 BY SIMILARITY.
FT DISULFID 265 334 BY SIMILARITY.
FT DISULFID 359 434 BY SIMILARITY.
FT DISULFID 391 407 BY SIMILARITY.
FT DISULFID 424 452 BY SIMILARITY.
FT CARBOHYD 185 185 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 403 403 N -> K (IN REF. 2).
FT CONFLICT 417 417 Y -> H (IN REF. 2).
FT CONFLICT 435 435 M -> R (IN REF. 2).
SQ SEQUENCE 477 AA; 53719 MW; 17486555C0E5077C CRC64;

Query Match 47.3%; Score 236; DB 1; Length 477;
Best Local Similarity 48.8%; Pred. No. 6.7e-20;
Matches 41; Conservative 13; Mismatches 30; Indels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASTDTMGRCPLPWNSATVLOQTYHAHRSNALQGLGKHNYCRNPDMR 61
DB 127 TCYKDGVTYRGTWSTSEGAQCINWNSMLLTRTYNGRRSDAITLGLGNHNYCRNPDMN 186
QY 62 RRPWCYVQVGLKPLVQECMVHDC 85
DB 187 SKPMCYVIKASKFILEFCSPVCS 210

RESULT 9
TPA_RAT
ID TPA_RAT STANDARD; PRT; 559 AA.
AC P19637;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE (t-PA) (t-plasminogen activator).

GN PLAT.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A. PubMed=3148445;
RX MEDLINE=89170114; PubMed=3148445;
RA Ny T., Leonardsson G., Hsueh A.J.W.;
RT "Cloning and characterization of a cDNA for rat tissue-type
RT plasminogen activator."
RL DNA 7:671-677(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90130448; PubMed=2105315;
RA Feng P., Ohlsson M., Ny T.;
RT "The structure of the TATA-less rat tissue-type plasminogen activator
RT gene. Species-specific sequence divergences in the promoter predict
RT differences in regulation of gene expression."
RL J. Biol. Chem. 265:2022-2027(1990).
CC -1- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZMOGEN PLASMINOGEN
CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CC CONTROLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
CC MANY OTHER PHYSIOPATHOLOGICAL EVENTS.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
CC BOND.
CC -1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC -1- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
CC -1- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CC CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; M23697; AAA41812.1; -;
DR EMBL; M31197; AAA42261.1; -;
DR EMBL; M31185; AAA42261.1; JOINED.
DR EMBL; M31186; AAA42261.1; JOINED.
DR EMBL; M31187; AAA42261.1; JOINED.
DR EMBL; M31188; AAA42261.1; JOINED.
DR EMBL; M31189; AAA42261.1; JOINED.
DR EMBL; M31190; AAA42261.1; JOINED.
DR EMBL; M31191; AAA42261.1; JOINED.
DR EMBL; M31192; AAA42261.1; JOINED.
DR EMBL; M31193; AAA42261.1; JOINED.
DR EMBL; M31194; AAA42261.1; JOINED.
DR EMBL; M31195; AAA42261.1; JOINED.
DR EMBL; M31196; AAA42261.1; JOINED.
DR EMBL; A19618; CAA01482.1; -;
DR PIR; A31597; A31597.
DR PIR; A35029; A35029.
DR HSSP; P00750; 1RTF.
DR MEROPS; S01.232; -;
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR000001; Kring1.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 1.

```

DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; fn1; 1.
DR SMART; SM00130; kr; 2.
DR SMART; SM00020; tryp_spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS50070; KRINGLE_2; 2.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Plasma; kringle; EGF-like domain; Repeat; Signal.
FT SIGNAL 1 17 PROBABLE.
FT PROPEP 18 29
FT CHAIN 30 559 TISSUE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 30 308 TISSUE-TYPE PLASMINOGEN ACTIVATOR A
FT CHAIN 309 559 TISSUE-TYPE PLASMINOGEN ACTIVATOR B
FT DOMAIN 36 78 FIBRONECTIN TYPE-I.
FT DOMAIN 79 117 EGF-LIKE.
FT DOMAIN 124 205 KRINGLE 1.
FT DOMAIN 203 294 KRINGLE 2.
FT DOMAIN 309 559 SERINE PROTEASE.
FT ACT_SITE 355 355 CHARGE RELAY SYSTEM.
FT ACT_SITE 404 404 CHARGE RELAY SYSTEM.
FT ACT_SITE 510 510 CHARGE RELAY SYSTEM.
FT DISULFID 38 68 BY SIMILARITY.
FT DISULFID 66 75 BY SIMILARITY.
FT DISULFID 83 94 BY SIMILARITY.
FT DISULFID 88 105 BY SIMILARITY.
FT DISULFID 107 116 BY SIMILARITY.
FT DISULFID 124 205 BY SIMILARITY.
FT DISULFID 145 187 BY SIMILARITY.
FT DISULFID 176 200 BY SIMILARITY.
FT DISULFID 213 294 BY SIMILARITY.
FT DISULFID 234 276 BY SIMILARITY.
FT DISULFID 265 289 BY SIMILARITY.
FT DISULFID 297 428 INTERCHAIN (BY SIMILARITY).
FT DISULFID 340 356 BY SIMILARITY.
FT DISULFID 348 417 BY SIMILARITY.
FT DISULFID 442 516 BY SIMILARITY.
FT DISULFID 474 490 BY SIMILARITY.
FT DISULFID 506 534 BY SIMILARITY.
FT CARBOHYD 149 149 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 481 481 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 380 380 E -> K (IN REF. 1).
SQ SEQUENCE 559 AA; 62903 MW; 7DBD3809C1D1C921 CRC64;

Query Match 45.3%; Score 226; DB 1; Length 559;
Best Local Similarity 49.4%; Pred. No. 1,1e-18;
Matches 41; Conservative 9; Mismatches 33; Indels 0; Gaps 0;

```

```

AC P49150;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator gamma precursor (EC 3.4.21.68) (DSPA
DE gamma).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OX NCBI_TaxID=9430;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=92039036; PubMed=1937019;
RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
RA Alagon A., Donner P., Schleuning W.D.;
RT "The plasminogen activator family from the salivary gland of the
RT vampire bat Desmodus rotundus: cloning and expression.";
RL Gene 105:229-237(1991).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=93393059; PubMed=1309059;
RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
RA Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common
RT vampire bat): unique fibrin specificity.";
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
CC -1- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
CC AGENT.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-1-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: MONOMER.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M63990; AAA31595.1; -.
DR HSSP; P98119; 1A5I.
DR MEROPS; S01.239; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; kringle; 1.
DR SMART; SM00130; kr; 1.
DR SMART; SM00020; tryp_spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW kringle; Signal; Multigene family.
FT SIGNAL 1 36 POTENTIAL.
FT CHAIN 37 394 SALIVARY PLASMINOGEN ACTIVATOR GAMMA.
FT DOMAIN 45 126 KRINGLE.
FT DOMAIN 142 394 SERINE PROTEASE.
FT ACT_SITE 189 189 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 238 238 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 345 345 CHARGE RELAY SYSTEM (BY SIMILARITY).

```



```
FT DISULFID 45 126 BY SIMILARITY.
FT DISULFID 66 108 BY SIMILARITY.
FT DISULFID 97 121 BY SIMILARITY.
FT DISULFID 131 262 BY SIMILARITY.
FT DISULFID 174 190 BY SIMILARITY.
FT DISULFID 182 251 BY SIMILARITY.
FT DISULFID 276 351 BY SIMILARITY.
FT DISULFID 308 324 BY SIMILARITY.
FT DISULFID 341 369 BY SIMILARITY.
FT CARBOHYD 315 315 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 394 AA; 44105 MW; 9CCD6F52F3D81FCD CRC64;

Query Match
Best Local Similarity 43.7%; Score 218; DB 1; Length 394;
Matches 38; Conservative 12; Mismatches 34; Indels 0; Gaps 0;

OY 2 TCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOTYHAHRSNALQLGKHNCRPNR 61
   |||: | ||| | | | | | | | | | | | | | | | | | | | | | | | |
Db 44 TCYKDGVTYRGTWSTSESGAQCINMNSNLLIRRTYNGRMPKLVKLGNGHNYCRNPDGA 103
   : ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 62 RRPWCYVQVGLKPLVQECMVHDC 85
   : ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 104 SKPWCYVTKARKFTSESCSVPCS 127
   : ||| | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 11
TPA_HUMAN STANDARD; PRT; 562 AA.
ID P00750; Q15103;
AC 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE (t-PA) (t-plasminogen activator) (Alteplase) (Retelase).
GN PLAT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Melanoma;
RX MEDLINE=83115262; PubMed=6337343;
RA Pennica D., Holmes W.E., Kohr W.J., Harkins R.N., Vohar G.A.,
RA Ward C.A., Bennett W.F., Yelverton E., Seeburg P.H., Heyneker H.L.,
RA Goeddel D.V., Collen D.;
RT "Cloning and expression of human tissue-type plasminogen activator
RT cDNA in E. coli.";
RL Nature 301:214-221(1983).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-Fetal lung;
RX MEDLINE=88262579; PubMed=3133640;
RA Sasaki H., Saito Y., Hayashi M., Otsuka K., Niwa M.;
RT "Nucleotide sequence of the tissue-type plasminogen activator cDNA
RT from human fetal lung cells.";
RL Nucleic Acids Res. 16:5695-5695(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=88054470; PubMed=2824147;
RA Reddy V.B., Garramone A.J., Sasak H., Wei C.-M., Watkins P., Galli J.,
RA Hsiung N.;
RT "Expression of human uterine tissue-type plasminogen activator in
RT mouse cells using BPV vectors.";
RL DNA 6:461-472(1987).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=86196143; PubMed=3009482;
RA Frielezer Degen S.J., Rajput B., Reich E.;
RT "The human tissue plasminogen activator gene.";
RL J. Biol. Chem. 261:6972-6985(1986).
RN [5]
RP SEQUENCE FROM N.A.
```

```
RX MEDLINE=84298137; PubMed=6089198;
RA Ny T., Elgh F., Lund B.;
RT "The structure of the human tissue-type plasminogen activator gene:
RT correlation of intron and exon structures to functional and
RT structural domains.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:5355-5359(1984).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=86284200; PubMed=3090401;
RA Harris T.J., Patel T., Marston F.A., Little S., Entage J.S.,
RA Opendakker G., Volckaert G., Rombaux W., Billiau A., Somer P.;
RT "Cloning of cDNA coding for human tissue-type plasminogen activator
RT and its expression in Escherichia coli.";
RL Mol. Biol. Med. 3:279-292(1986).
RN [7]
RP SEQUENCE OF 212-361 FROM N.A.
RX MEDLINE=83169656; PubMed=6572897;
RA Edlund T., Ny T., Raanby M., Heden L.-O., Palm G., Holmgren E.,
RA Josephson S.;
RT "Isolation of cDNA sequences coding for a part of human tissue
RT plasminogen activator.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:349-352(1983).
RN [8]
RP SEQUENCE OF 1-36 FROM N.A.
RX MEDLINE=85289338; PubMed=3161893;
RA Fisher R., Waller E.K., Grossi G., Thompson D., Tizard R.,
RA Schleuning W.-D.;
RT "Isolation and characterization of the human tissue-type plasminogen
RT activator structural gene including its 5' flanking region.";
RL J. Biol. Chem. 260:11223-11230(1985).
RN [9]
RP SEQUENCE OF 31-562 FROM N.A.
RX MEDLINE=91291340; PubMed=1368681;
RA Itagaki Y., Yasuda H., Morinaga T., Mitsuda S., Higashio K.;
RT "Purification and characterization of tissue plasminogen activator
RT secreted by human embryonic lung diploid fibroblasts, IMR-90 cells.";
RL Agric. Biol. Chem. 55:1225-1232(1991).
RN [10]
RP SEQUENCE OF 36-562.
RC TISSUE-Melanoma;
RX MEDLINE=85000468; PubMed=6433976;
RA Pohl G., Kaellstroem M., Bergsdorf N., Wallen P., Joernvall H.;
RT "Tissue plasminogen activator: peptide analyses confirm an indirectly
RT derived amino acid sequence, identify the active site serine residue,
RT establish glycosylation sites, and localize variant differences.";
RL Biochemistry 23:3701-3707(1984).
RN [11]
RP SEQUENCE OF 33-52 AND 311-330.
RC TISSUE-Melanoma;
RX MEDLINE=83209620; PubMed=6682760;
RA Wallen P., Pohl G., Bergsdorf N., Raanby M., Ny T., Joernvall H.;
RT "Purification and characterization of a melanoma cell plasminogen
RT activator.";
RL Eur. J. Biochem. 132:681-686(1983).
RN [12]
RP SEQUENCE FROM N.A. (SHORT ISOFORM).
RC TISSUE-Umbilical vein;
RX MEDLINE=90192129; PubMed=2107528;
RA Siebert P.D., Fong K.;
RT "Variant tissue-type plasminogen activator (PLAT) cDNA obtained from
RT human endothelial cells.";
RL Nucleic Acids Res. 18:1086-1086(1990).
RN [13]
RP STRUCTURE OF CARBOHYDRATES.
RX MEDLINE=90092112; PubMed=2513186;
RA Pfeiffer G., Schmidt M., Strube K.-H., Geyer R.;
RT "Carbohydrate structure of recombinant human uterine tissue
RT plasminogen activator expressed in mouse epithelial cells.";
RL Eur. J. Biochem. 186:273-286(1989).
RN [14]
RP CARBOHYDRATE-LINKAGE SITE THR-96.
RX MEDLINE=91159408; PubMed=1900431;
RA Harris R.J., Leonard C.K., Guzzetta A.W., Spellman M.W.;
```

RT "Tissue plasminogen activator has an O-linked fucose attached to
 RT threonine-61 in the epidermal growth factor domain."
 RL Biochemistry 30:2311-2314(1991).
 RP [15]
 RP DISULFIDE BONDS IN KRINGLE 2.
 RX MEDLINE=91244765; PubMed=1645336;
 RA Vlahos C.J., Wilhelm O.G., Hassell T., Taskunas S.R., Bang N.U.;
 RT "Disulfide pairing of the recombinant kringle-2 domain of tissue
 RT plasminogen activator produced in *Escherichia coli*."
 RL J. Biol. Chem. 266:10070-10072(1991).
 RN [16]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF CATALYTIC DOMAIN.
 RX MEDLINE=96200985; PubMed=8613982;
 RA Lamba D., Bauer M., Huber R., Fischer S., Rudolph R., Kohnert U.,
 RA Bode W.;
 RT "The 2.3 A crystal structure of the catalytic domain of recombinant
 RT two-chain human tissue-type plasminogen activator."
 RL J. Mol. Biol. 258:117-135(1996).
 RN [17]
 RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF CATALYTIC DOMAIN.
 RX MEDLINE=97449126; PubMed=9305622;
 RA Renatus M., Engh R.A., Stubbs M.T., Huber R., Fischer S., Kohnert U.,
 RA Bode W.;
 RT "Lysine 156 promotes the anomalous proenzyme activity of tPA: X-ray
 RT crystal structure of single-chain human tPA."
 RL EMBO J. 16:4797-4805(1997).
 RN [18]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF KRINGLE 2.
 RX MEDLINE=92118803; PubMed=1310033;
 RA de Vos A., Ultsch M.H., Kelley R.F., Padmanabhan K., Tulinskly A.,
 RA Westbrook M.L., Kossiakof A.A.;
 RT "Crystal structure of the kringle 2 domain of tissue plasminogen
 RT activator at 2.4-A resolution."
 RL Biochemistry 31:270-279(1992).
 RN [19]
 RP STRUCTURE BY NMR OF KRINGLE 2.
 RX MEDLINE=90122799; PubMed=2558718;
 RA Byeon I.-J.L., Kelley R.F., Llinas M.;
 RT "1H NMR structural characterization of a recombinant kringle 2 domain
 RT from human tissue-type plasminogen activator."
 RL Biochemistry 28:9350-9360(1989).
 RN [20]
 RP STRUCTURE BY NMR OF KRINGLE 2.
 RX MEDLINE=91200042; PubMed=1901789;
 RA Byeon I.-J.L., Kelley R.F., Llinas M.;
 RT "Kringle-2 domain of the tissue-type plasminogen activator. 1H-NMR
 RT assignments and secondary structure."
 RL Eur. J. Biochem. 197:155-165(1991).
 RN [21]
 RP STRUCTURE BY NMR OF KRINGLE 2.
 RX MEDLINE=92106329; PubMed=1762144;
 RA Byeon I.-J.L., Llinas M.;
 RT "Solution structure of the tissue-type plasminogen activator kringle
 RT 2 domain complexed to 6-aminohexanoic acid an antifibrinolytic
 RT drug."
 RL J. Mol. Biol. 222:1035-1051(1991).
 RN [22]
 RP STRUCTURE BY NMR OF 38-85.
 RX MEDLINE=92292163; PubMed=1602484;
 RA Downing A.K., Driscoll P.C., Harvey T.S., Dudgeon T.J., Smith B.O.,
 RA Baron M., Campbell I.D.;
 RT "Solution structure of the fibrin binding finger domain of
 RT tissue-type plasminogen activator determined by 1H nuclear magnetic
 RT resonance."
 RL J. Mol. Biol. 225:821-833(1992).
 RN [23]
 RP STRUCTURE BY NMR OF 36-126.
 RX MEDLINE=96027104; PubMed=7582899;
 RA Smith B.O., Downing A.K., Driscoll P.C., Dudgeon T.J., Campbell I.D.;
 RT "The solution structure and backbone dynamics of the fibronectin type
 RT I and epidermal growth factor-like pair of modules of tissue-type
 RT plasminogen activator."
 RL Structure 3:823-833(1995).

CC -I- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZMOGEN PLASMINOGEN
 CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
 CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
 CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
 CC MANY OTHER PHYSIOLOGICAL EVENTS.
 CC -I- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -I- SUBUNIT: BINDS TO FIBRIN WITH HIGH AFFINITY. THIS INTERACTION
 CC
 CC Query Match 43.3%; Score 216; DB 1; Length 562;
 CC Best Local Similarity 47.0%; Pred. No. 1.6e-17;
 CC Matches 39; Conservative 9; Mismatches 35; Indels 0; Gaps 0;
 QY 3 CYEGNGHFYRGKASTDTMGRPCLEPWN SATVLQQT YHAHRSNALQLGKHNCRPNRR 62
 DB 215 CYFGNSAYRGTHSLTESGASCLPWN SMILIGKYTAQNP SAQALGLGKHNCRPNR 274
 QY 63 RPYCYVQVGLKPLVQECMVHDA 85
 DB 275 KPWCHVLKNNRLTWECYDVPSCS 297
 RESULT 12
 TPA_MOUSE STANDARD; PRT; 559 AA.
 ID TPA_MOUSE
 AC P11214;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
 DE (t-PA) (t-plasminogen activator).
 GN PLAT.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88087303; PubMed=2826484;
 RA Rickles R.J., Darrow A.L., Strickland S.;
 RT "Molecular cloning of complementary DNA to mouse tissue plasminogen
 RT activator mRNA and its expression during F9 teratocarcinoma cell
 RT differentiation."
 RL J. Biol. Chem. 263:1563-1569(1988).
 CC -I- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZMOGEN PLASMINOGEN
 CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
 CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
 CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
 CC MANY OTHER PHYSIOLOGICAL EVENTS.
 CC -I- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -I- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
 CC BOND.
 CC -I- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
 CC -I- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
 CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
 CC ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
 CC -I- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
 CC CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
 CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -I- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
 CC -I- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL; J03520; AAA40470.1; -.

DR PIR; A29941; A29941.
DR HSP; P00750; 1A5H.
DR MEROPS; S01.232; -.
DR MGD; MGI:97610; Plat.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR000001; Kringl.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fnl; 1.
DR Pfam; PF00051; kringl; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringl; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FNI; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS50070; KRINGLE_2; 2.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
FT SIGNAL; 1 17
FT PROPEP; 18 29
FT CHAIN; 30 308
FT CHAIN; 309 559
FT CHAIN; 309 559
FT DOMAIN; 36 78
FT DOMAIN; 79 117
FT DOMAIN; 124 205
FT DOMAIN; 213 294
FT DOMAIN; 309 559
FT ACT_SITE; 355 355
FT ACT_SITE; 404 404
FT ACT_SITE; 510 510
FT DISULFID; 38 68
FT DISULFID; 66 75
FT DISULFID; 83 94
FT DISULFID; 88 105
FT DISULFID; 107 116
FT DISULFID; 124 205
FT DISULFID; 145 187
FT DISULFID; 176 200
FT DISULFID; 213 294
FT DISULFID; 234 276
FT DISULFID; 265 289
FT DISULFID; 297 428
FT DISULFID; 340 356
FT DISULFID; 348 417
FT DISULFID; 442 516
FT DISULFID; 474 490
FT DISULFID; 506 534
FT CARBOHYD; 149 149
FT CARBOHYD; 481 481
SQ SEQUENCE 559 AA; 63110 MW; 4ACE57DC6A282A5 CRC64;

Query Match 43.1%; Score 215; DB 1; Length 559;
Best Local Similarity 47.0%; Pred. No. 2.1e-17;
Matches 39; Conservative 10; Mismatches 34; Indels 0; Gaps 0;

OY 2 TCYEGNGHFYRGAATDTMGRPCLPWNATVLOOTYHAHRSNALOLGLGKHNYPNDNR 61
Db 123 TCFEEGGITRTGTWSTAESGAECINWSSVSLKPYNARRPNAIKLGLGNHNYCRNPDRD 182

OY 62 RRPWCYVGVGLKPLVQECMVHDC 84
Db 183 LKFWCYVFKAGKYTTFCSTPAC 205

RESULT 13
TPA_BOVIN
ID TPA_BOVIN STANDARD; PRT; 566 AA.
AC Q28198;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (TPA)
DE (t-PA) (t-plasminogen activator).
GN PLAT.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Ravn P., Berglund L., Petersen T.E.;
RT "Cloning and characterization of the bovine plasminogen activators uPA
and tPA."
RT Int. Dairy J. 5:605-617(1995).
CC -1- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZMOGEN PLASMINOGEN
TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
MANY OTHER PHYSIOLOGICAL EVENTS.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
plasminogen to form plasmin.
CC -1- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
BOND.
CC -1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC -1- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
ARG-314 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
CC -1- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; X85800; CAA59795.1; -.
DR HSP; P00750; 1RTE.
DR MEROPS; S01.232; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR000001; Kringl.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fnl; 1.
DR Pfam; PF00051; kringl; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringl; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FNI; 1.

```
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolyase; Serine protease; Glycoprotein;
KW Plasma; Kringle; EGF-like domain; Repeat; Signal.
FT SIGNAL 1 21 BY SIMILARITY.
FT PROPEP 22 33 BY SIMILARITY.
FT CHAIN 34 566 TISSUE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 34 314 TISSUE-TYPE PLASMINOGEN ACTIVATOR A
FT CHAIN 315 566 TISSUE-TYPE PLASMINOGEN ACTIVATOR B
FT DOMAIN 40 82 CHAIN.
FT DOMAIN 83 121 FIBRONECTIN TYPE-I.
FT DOMAIN 128 209 EGF-LIKE.
FT DOMAIN 219 300 KRINGLE 1.
FT DOMAIN 315 566 KRINGLE 2.
FT ACT_SITE 361 410 SERINE PROTEASE.
FT ACT_SITE 410 410 CHARGE RELAY SYSTEM.
FT ACT_SITE 517 517 CHARGE RELAY SYSTEM.
FT DISULFID 42 72 BY SIMILARITY.
FT DISULFID 70 79 BY SIMILARITY.
FT DISULFID 87 98 BY SIMILARITY.
FT DISULFID 92 109 BY SIMILARITY.
FT DISULFID 111 120 BY SIMILARITY.
FT DISULFID 128 209 BY SIMILARITY.
FT DISULFID 149 191 BY SIMILARITY.
FT DISULFID 180 204 BY SIMILARITY.
FT DISULFID 219 300 BY SIMILARITY.
FT DISULFID 240 282 BY SIMILARITY.
FT DISULFID 271 295 BY SIMILARITY.
FT DISULFID 303 434 INTERCHAIN (BY SIMILARITY).
FT DISULFID 346 362 BY SIMILARITY.
FT DISULFID 354 423 BY SIMILARITY.
FT DISULFID 448 523 BY SIMILARITY.
FT DISULFID 480 496 BY SIMILARITY.
FT DISULFID 513 541 BY SIMILARITY.
FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 487 487 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 566 AA; 63701 MW; 2EB6BBE4E32276C3 CRC64;

Query Match 42.9%; Score 214; DB 1; Length 566;
Best Local Similarity 46.4%; Pred. No. 2.8e-17;
Matches 39; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHFYRGRASTDTMGRPCLPWNATVLOQTYHAHRSNALQGLGKHNCRPNDR 61
   |||: | ||| | | | | | | | | | | | | | | | | | | | | | | | |
DB 127 TCYKDGVAAYRGVSTAESGAECANWNSSGLAMKPYSGRRPNAIRLGLGNHNYCRNPQD 186

QY 62 RRPWCYVQVGLKPLVQECMVHDC 85
   :||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 187 SKPMCYVFKAQKXISEFCSTPACA 210

RESULT 14
URTI_DESRO
ID URTI_DESRO STANDARD; PRT; 477 AA.
AC P98119;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator alpha 1 precursor (EC 3.4.21.68) (DSPA
DE alpha-1).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
```

```
OC Desmodontinae; Desmodus.
OX NCBI_TaxID=9430;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=92039036; PubMed=1937019;
RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
RT Alagon A., Donner P., Schleuning W.D.;
RT "The plasminogen activator family from the salivary gland of the
RT vampire bat Desmodus rotundus: cloning and expression.";
RL Gene 105:229-237(1991).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=93393059; PubMed=1309059;
RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
RA Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common
RT vampire bat): unique fibrin specificity.";
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RC TISSUE=Salivary gland;
RX MEDLINE=98022741; PubMed=9354616;
RA Renatus M., Stubbs M.T., Huber R., Bringmann P., Donner P.,
RA Schleuning W.D., Bode W.;
RT "Catalytic domain structure of vampire bat plasminogen activator: a
RT molecular paradigm for proteolysis without activation cleavage.";
RL Biochemistry 36:13483-13493(1997).
CC -!- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. POTENT THROMBOLYTIC AGENT.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- ENZYME REGULATION: ACTIVITY TOWARD PLASMINOGEN IS STIMULATED IN
CC THE PRESENCE OF FIBRIN I.
CC -!- SUBUNIT: MONOMER.
CC -!- DOMAIN: THE FIBRONECTIN TYPE-I DOMAIN MEDIATES BINDING TO FIBRIN,
CC AND THE KRINGLE DOMAIN APPARENTLY MEDIATES FIBRIN-INDUCED
CC STIMULATION OF ACTIVITY.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M63987; AAA31591.1; -.
DR EMBL; M63986; AAA31592.1; -.
DR PDB; 1A51; 23-MAR-99.
DR MEROPS; S01.232; -.
DR GlycoSuiteDB; P98119; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
```


DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasmalogen activation; Hydrolase; Serine protease; Glycoprotein;
Kringles; EGF-like domain; Signal; Multigene family; 3D-structure.
FT SIGNAL; 1 36
FT CHAIN 37 477
FT DOMAIN 40 82
FT DOMAIN 83 121
FT DOMAIN 128 209
FT DOMAIN 225 477
FT ACT_SITE 272 321
FT ACT_SITE 321 428
FT ACT_SITE 428 428
FT DISULFID 42 72
FT DISULFID 70 79
FT DISULFID 87 98
FT DISULFID 92 109
FT DISULFID 111 120
FT DISULFID 128 209
FT DISULFID 149 191
FT DISULFID 180 204
FT DISULFID 214 345
FT DISULFID 257 273
FT DISULFID 265 334
FT DISULFID 359 434
FT DISULFID 391 407
FT DISULFID 424 452
FT CARBOHYD 153 153
FT CARBOHYD 398 398
FT SEQUENCE 477 AA; 53616 MW; AA06FD1739C10E5E CRC64;
SQ
Query Match 41.7%; Score 208; DB 1; Length 477;
Best Local Similarity 45.2%; Pred. No. 1.2e-16;
Matches 38; Conservative 11; Mismatches 35; Indels 0; Gaps 0;
QY 2 TCYEGNGHFYRGKASTDTMGRPCLPWNSATVLO-QTYHAHRSNALQLGLGKHNYCRNPDNR 61
DB 127 TCYEGOGVTYRGTWSTAESRYECINWNSSLLTRTYNGRMPDAFNGLGNHNYCRNPNCA 186
QY 62 RRPWCYVQGLKPLVQECMVHDC 85
DB 187 RRPWCYVIKAGFTSESCSVPC 210
RESULT 15
UROC_CHICK STANDARD; PRT; 434 AA.
ID UROC_CHICK
AC P15120;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (UPA)
DE (U-plasminogen activator).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90110185; PubMed=2295632;
RA Leslie N.D., Kessler C.A., Bell S.M., Degen J.L.;
RT "The chicken urokinase-type plasminogen activator gene."
RL J. Biol. Chem. 265:1339-1344(1990).
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in

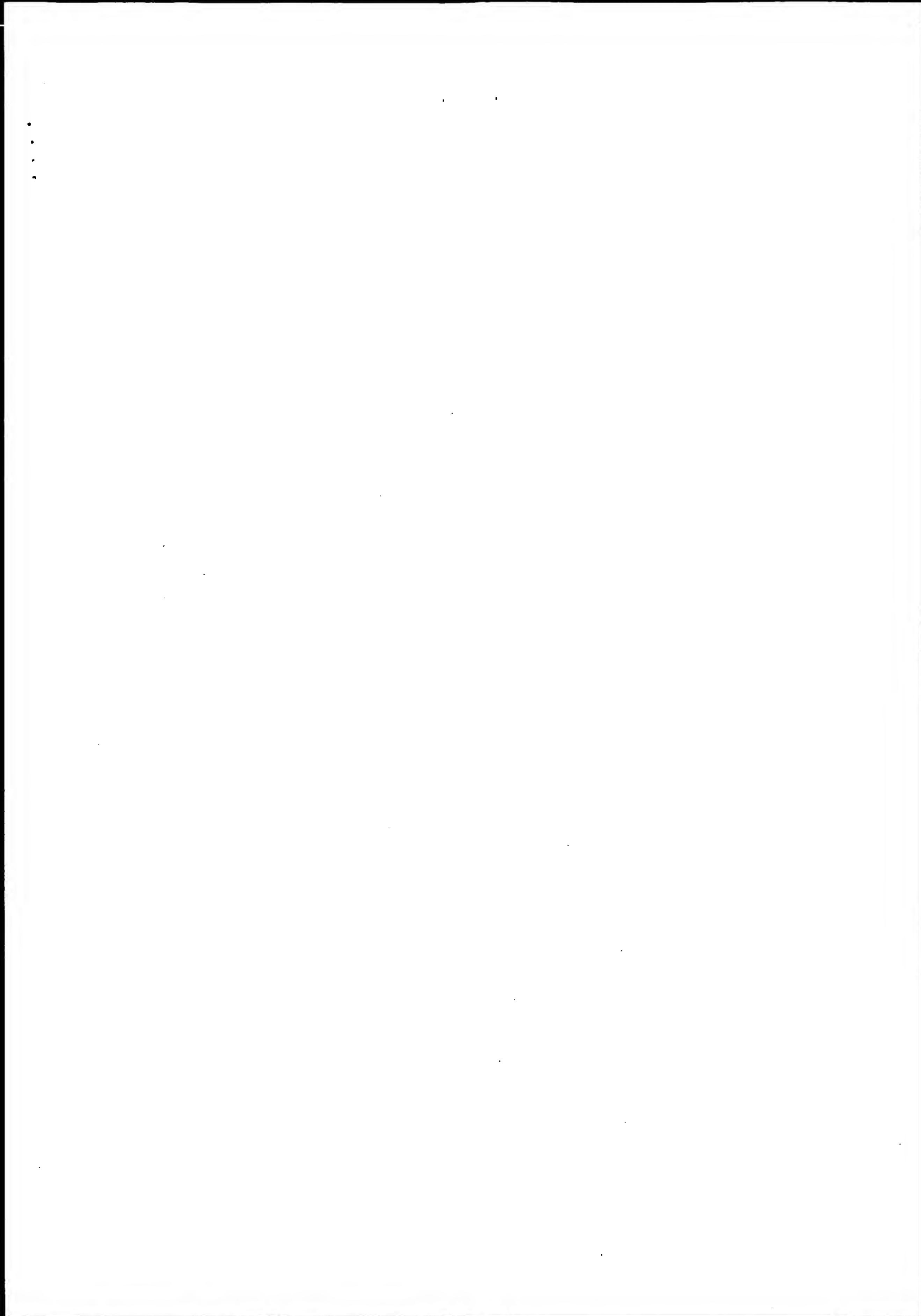
CC plasmalogen to form plasmin.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J05187; AAA49131.1; -
CC EMBL; J05188; AAA49130.1; -
CC PIR; A35005; A35005.
CC HSSP; P00763; IDPO.
CC MEROPS; S01.231; -
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR000561; EGF-like.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Ser-protease_Try.
CC Pfam; PF00051; Kringles; 1.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC ProDom; PD000395; Kringle; 1.
CC SMART; SM00181; EGF; 1.
CC SMART; SM00130; KR; 1.
CC SMART; SM00020; Tryp_Spc; 1.
CC PROSITE; PS00022; EGF_1; 1.
CC PROSITE; PS01186; EGF_2; 1.
CC PROSITE; PS00021; KRINGLE_1; 1.
CC PROSITE; PS50070; KRINGLE_2; 1.
CC PROSITE; PS50240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasmalogen activation; Hydrolase; Serine protease; Glycoprotein;
Kringles; EGF-like domain; Signal; zymogen.
FT SIGNAL 1 20
FT CHAIN 21 434
FT CHAIN 21 171
FT CHAIN 173 434
FT CHAIN 36 72
FT DOMAIN 79 158
FT DOMAIN 159 172
FT DOMAIN 173 434
FT DISULFID 40 48
FT DISULFID 42 60
FT DISULFID 62 71
FT DISULFID 162 296
FT DISULFID 202 218
FT DISULFID 210 285
FT DISULFID 310 379
FT DISULFID 342 358
FT DISULFID 369 397
FT ACT_SITE 217 217
FT ACT_SITE 272 272
FT ACT_SITE 373 373
FT CARBOHYD 228 228
FT SEQUENCE 434 AA; 49400 MW; BD881048DD66A55 CRC64;
SQ
Query Match 40.8%; Score 203.5; DB 1; Length 434;
Best Local Similarity 55.9%; Pred. No. 3.5e-16;
Matches 38; Conservative 6; Mismatches 19; Indels 5; Gaps 2;
QY 3 CYEGNGHFYRGKASTDTMGRPCLPWNSATVLO-QTYHAHRSNALQLGLGKHNYCRNPDNR 61
DB 79 CYSGNGEDYRGMADP-----GCLYWDHPSVIRWGDYHADLKNALQLGLGKHNYCRNPNCR 134
QY 62 RRPWCYVQ 69
DB 135 SRPWCYTK 142

Tue Nov 12 18:06:47 2002

pct-us02-27855-1.rsp

Page 15

Search completed: November 11, 2002, 13:00:25
Job time : 22 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2002, 12:24:41 ; Search time 75 Seconds
(without alignments)
236.267 Million cell updates/sec

Title: PCT-US02-27855-1
Perfect score: 499
Sequence: 1 KTCYEGNGHFRYRKASTDTM.....YVQVGLKPLVQECMVHDCAD 86

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriophage:*
- 17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	484	97.0	154	4 Q96SE8	Q96SE8 homo sapien
2	383	76.8	157	6 Q9TVA8	Q9TVA8 bos taurus
3	306	61.3	214	6 Q9XT70	Q9XT70 oryctolagus
4	217	43.5	559	11 Q91VP2	Q91VP2 mus musculus
5	216	43.3	395	4 Q9BZW1	Q9BZW1 homo sapien
6	216	43.3	516	4 Q9BU99	Q9BU99 homo sapien
7	211	42.3	562	6 Q8SQ23	Q8SQ23 sus scrofa
8	191	38.3	202	13 Q90675	Q90675 gallus gall
9	188.5	37.8	560	4 Q14520	Q14520 homo sapien
10	188	37.7	653	11 Q8VCS4	Q8VCS4 mus musculus
11	186.5	37.4	597	11 Q35727	Q35727 sus scrofa
12	177.5	35.6	616	6 Q97507	Q97507 sus scrofa
13	157	31.5	452	13 Q90Y90	Q90Y90 xenopus lae
14	153.5	30.8	812	11 Q9R0W3	Q9R0W3 rattus norv
15	153	30.7	454	6 Q46506	Q46506 papio hamad
16	153	30.7	806	6 O18783	O18783 macropus eu

17	152	30.5	113	4 Q9UIR6	Q9UIR6 homo sapien
18	151.5	30.4	359	6 Q8WMR1	Q8WMR1 canis fam11
19	151	30.3	113	4 Q9UIR7	Q9UIR7 homo sapien
20	151	30.3	113	4 Q9UIR5	Q9UIR5 homo sapien
21	149	29.9	810	4 Q15146	Q15146 homo sapien
22	148.5	29.8	709	13 Q90ZN6	Q90ZN6 brachydanio
23	147.5	29.6	399	4 Q96GL8	Q96GL8 homo sapien
24	147.5	29.6	420	4 Q9BTP9	Q9BTP9 homo sapien
25	147.5	29.6	716	13 Q91691	Q91691 xenopus lae
26	147	29.5	393	4 Q9BRB6	Q9BRB6 homo sapien
27	146	29.3	105	4 Q9UIR8	Q9UIR8 homo sapien
28	145	29.1	648	4 Q9HIV4	Q9HIV4 homo sapien
29	144	28.9	567	4 Q13208	Q13208 homo sapien
30	142	28.5	902	5 Q17576	Q17576 caenorhabdi
31	142	28.5	928	5 Q9BLV1	Q9BLV1 caenorhabdi
32	139.5	28.0	717	13 P70006	P70006 xenopus lae
33	138	27.7	1145	5 Q9BRL8	Q9BRL8 aplysia cal
34	137.5	27.6	704	13 Q90865	Q90865 gallus gall
35	137	27.5	812	11 Q91WJ5	Q91WJ5 mus musculus
36	135	27.1	132	4 Q16609	Q16609 homo sapien
37	135	27.1	145	6 Q28911	Q28911 macaca fasc
38	134	26.9	685	5 Q24488	Q24488 drosophila
39	133.5	26.8	313	13 Q9PU78	Q9PU78 crocodylus
40	132	26.5	263	4 Q00318	Q00318 homo sapien
41	132	26.5	263	4 Q96FE7	Q96FE7 homo sapien
42	131.5	26.4	215	13 Q42341	Q42341 gallus gall
43	131.5	26.4	716	11 Q91XG8	Q91XG8 mus musculus
44	131.5	26.4	726	13 Q90978	Q90978 gallus gall
45	131.5	26.4	2869	6 Q28398	Q28398 erinaceus e

ALIGNMENTS

RESULT 1

ID	Q96SE8	PRELIMINARY:	PRT:	154 AA.
AC	Q96SE8;			
DT	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)			
DE	Urokinase-type plasminogen activator amino-terminal fragment.			
GN	ATP.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Fu J., Bai X., Ruan C.;			
RT	"Cloning and expression of the amino-terminal fragment of human			
RT	urokinase-type plasminogen activator.";			
RL	Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Bai X., Wang W., Xi X., Ruan C.;			
RT	"Overexpression of the amino-terminal fragment of human urokinase-type			
RT	plasminogen activator in breast cancer cells results in decreased			
RL	tumor invasion, growth and angiogenesis.";			
RL	Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AY029537; AAK38734.1; -			
DR	InterPro; IPR000561; EGF-like.			
DR	InterPro; IPR000001; Kringle.			
DR	Pfam; PF00051; Kringle; 1.			
DR	ProDom; PD000395; Kringle; 1.			
DR	SMART; SM00181; EGF; 1.			
DR	PROSITE; PS00022; EGF_1; UNKNOWN_1.			
DR	PROSITE; PS00021; KRINGLE_1; UNKNOWN_1.			
DR	PROSITE; PS50070; KRINGLE_2; 1.			
KW	Kinase.			
SQ	SEQUENCE 154 AA; 17305 MW; A3CCF2FCFF505572 CRC64;			

Query Match 97.0%; Score 484; DB 4; Length 154;
Best Local Similarity 97.7%; Pred. No. 3.2e-52;
Matches 84; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWNASATVLOQTYHAHRSNALQLGKHNCRPN 60
DB 68 KTCYEGNGHFYRGKASTDTMGRCPLPWNASATVLOQTYHAHRSNALQLGKHNCRPN 127
QY 61 RRRPWCYVOVGLKPLVQECMVHDCAD 86
DB 128 RRRPWCYVOVGLKPLVQECMVHDCAD 153

RESULT 2

Q9TVAB PRELIMINARY; PRT; 157 AA.
AC Q9TVAB; PRT; 157 AA.
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Urokinase plasminogen activator (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_Taxid=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-SKELETAL MUSCLE;
RX MEDLINE=21071388; PubMed=11204721;
RA Balcerzak D., Querengueser L., Dixon W.T., Baracos V.E.;
RT "Coordinate expression of matrix-degrading proteinases and their
activators and inhibitors in bovine skeletal muscle.";
RL J. Anim. Sci. 79:94-107(2001).
DR EMBL; AF144761; AAD30301.1; -.
DR HSSP; P00749; IURK.
DR MEROPS; S01.231; -.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
KW Kinase.
FT NON_TER
FT NON_TER
SQ SEQUENCE 157 AA; 17858 MW; A768D6C72C1FBFB7 CRC64;

Query Match 76.8%; Score 383; DB 6; Length 157;
Best Local Similarity 75.3%; Pred. No. 1.1e-39;
Matches 64; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWNASATVLOQTYHAHRSNALQLGKHNCRPN 60
DB 36 KTCYEGNGHFYRGKASTDTMGRCPLPWNASATVLOQTYHAHRSNALQLGKHNCRPN 95
QY 61 RRRPWCYVOVGLKPLVQECMVHDCAD 85
DB 96 RRRPWCYVOVGLKPLVQECMVHDCAD 120

RESULT 3

Q9XT70 PRELIMINARY; PRT; 214 AA.
AC Q9XT70; PRT; 214 AA.
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Urokinase-type plasminogen activator (Fragment).
OS Oryctolagus cuniculus (Rabbit).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_Taxid=9986;

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-LUNG;
RA Yin J., Idell S.;
RT "Partial mRNA of rabbit uPA.";
Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
TRYPSIN FAMILY.
CC EMBL; AF097647; AAD39351.1; -.
DR HSSP; P00749; IJUN.
DR MEROPS; S01.231; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser-protease_Try.
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
KW Hydrolyase; Kinase; Serine protease.
FT NON_TER
FT NON_TER
SQ SEQUENCE 214 AA; 24314 MW; 69975C41C32B0D7E CRC64;

Query Match 61.3%; Score 306; DB 6; Length 214;
Best Local Similarity 81.2%; Pred. No. 5.6e-30;
Matches 52; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 20 MGRPCLPWNASATVLOQTYHAHRSNALQLGKHNCRPNRRPWCYVOVGLKPLVQEC 79
DB 1 MDRPCLPWNASATVLOQTYHAHRSNALQLGKHNCRPNRRPWCYVOVGLKPLVQEC 60
QY 80 MVHD 83
DB 61 KVHD 64

RESULT 4

Q91VP2 PRELIMINARY; PRT; 559 AA.
AC Q91VP2; PRT; 559 AA.
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Similar to plasminogen activator, tissue.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BREAST TUMOR;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC011256; AAH1256.1; -.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000083; Fibronctn.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser-protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; Kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.

DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS01253; FIBRONECTIN_1; UNKNOWN_1.
DR PROSITE; PS00021; KRINGLE_1; UNKNOWN_2.
DR PROSITE; PS50070; KRINGLE_2; 2.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; UNKNOWN_1.
KW Hydrolyase; Serine protease.
SQ SEQUENCE 559 AA; 63122 MW; 8CCEE2BD894514D9 CRC64;

Query Match 43.5%; Score 217; DB 11; Length 559;
Best Local Similarity 47.0%; Pred. No. 1.8e-18;
Matches 39; Conservative 11; Mismatches 33; Indels 0; Gaps 0;

QY 3 CYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGLGKHNYCRNPDNR 62
DB 213 CYVGKGVYTRGTHSLTSGASCLPWNISVLGMKSYTAMRINSQALGLGRHNYCRNPDGDA 272
QY 63 RPWCYVQVGLKPLVQECMVHDC 85
DB 273 RPWCHVMKDKRLTWEYCDVPCS 295

RESULT 5

Q9BZW1 PRELIMINARY; PRT; 395 AA.
AC Q9BZW1;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Neonatal thrombolytic agent alpha-form (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;

RN [1]
RP SEQUENCE FROM N.A.

RA Dou D.;

RT "A brain-type plasminogen activator";
RL Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE

CC TRYPSIN FAMILY.

DR EMBL; AF260825; AAK11956.1; -.

DR HSSP; P00750; 1PK2.

DR MEROPS; S01.232; -.

DR InterPro; IPR001314; Chymotrypsin.

DR InterPro; IPR000083; Fibrinctn1.

DR InterPro; IPR000001; Kringle.

DR InterPro; IPR001254; Ser_protease_Try.

DR Pfam; PF00039; fn1; 1.

DR Pfam; PF00051; kringle; 1.

DR Pfam; PF00089; trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR PRINTS; PR00018; KRINGLE.

DR ProDom; PD000395; kringle; 1.

DR SMART; SM00058; FN1; 1.

DR SMART; SM00130; KR; 1.

DR SMART; SM00020; TRYP_SPE; 1.

DR PROSITE; PS01253; FIBRONECTIN_1; 1.

DR PROSITE; PS00021; KRINGLE_1; UNKNOWN_1.

DR PROSITE; PS50070; KRINGLE_2; 1.

DR PROSITE; PS50240; TRYPSIN_DOM; 1.

DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.

DR PROSITE; PS00135; TRYPSIN_SER; 1.

KW Hydrolyase; Serine protease.

FT NON_TER 395
SQ SEQUENCE 395 AA; 44323 MW; 3FBD4A2F0B7C11C8 CRC64;

Query Match 43.3%; Score 216; DB 4; Length 395;
Best Local Similarity 47.0%; Pred. No. 1.6e-18;
Matches 39; Conservative 9; Mismatches 35; Indels 0; Gaps 0;

QY 3 CYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGLGKHNYCRNPDNR 62

DB 48 CYFGNGSAYRGTHTSLTESGASCLPWNISMILLIGKYTAQNPSAQLGLGKHNYCRNPDGDA 107
QY 63 RPWCYVQVGLKPLVQECMVHDC 85
DB 108 KPWCYVGLKPLVQECMVHDC 85

RESULT 6

Q9BU99 PRELIMINARY; PRT; 516 AA.

AC Q9BU99;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Similar to plasminogen activator, tissue.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;

RN [1]
RP SEQUENCE FROM N.A.

RC TISSUE=SKIN;

RA Strausberg R.;

RL Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE

CC TRYPSIN FAMILY.

DR EMBL; BC002795; AAH02795.1; -.

DR HSSP; P00750; 1A5H.

DR MEROPS; S01.232; -.

DR InterPro; IPR001314; Chymotrypsin.

DR InterPro; IPR000561; EGF-like.

DR InterPro; IPR000001; Kringle.

DR InterPro; IPR001254; Ser_protease_Try.

DR Pfam; PF00008; EGF; 1.

DR Pfam; PF00051; kringle; 2.

DR Pfam; PF00089; trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR PRINTS; PR00018; KRINGLE.

DR ProDom; PD000395; Kringle; 2.

DR SMART; SM00181; EGF; 1.

DR SMART; SM00001; EGF_like; 1.

DR SMART; SM00130; KR; 2.

DR SMART; SM00020; TRYP_SPE; 1.

DR PROSITE; PS00022; EGF_1; UNKNOWN_1.

DR PROSITE; PS01186; EGF_2; 1.

DR PROSITE; PS00021; KRINGLE_1; UNKNOWN_2.

DR PROSITE; PS50070; KRINGLE_2; 2.

DR PROSITE; PS50240; TRYPSIN_DOM; 1.

DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.

DR PROSITE; PS00135; TRYPSIN_SER; 1.

KW EGF-like domain; Glycoprotein; Hydrolyase; Serine protease.

SQ SEQUENCE 516 AA; 57370 MW; BAB31901FDC96800 CRC64;

Query Match 43.3%; Score 216; DB 4; Length 516;
Best Local Similarity 47.0%; Pred. No. 2.2e-18;
Matches 39; Conservative 9; Mismatches 35; Indels 0; Gaps 0;

QY 3 CYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGLGKHNYCRNPDNR 62
DB 169 CYFGNGSAYRGTHTSLTESGASCLPWNISMILLIGKYTAQNPSAQLGLGKHNYCRNPDGDA 228
QY 63 RPWCYVQVGLKPLVQECMVHDC 85
DB 229 KPWCYVGLKPLVQECMVHDC 85

RESULT 7

Q8SQ23 PRELIMINARY; PRT; 562 AA.

AC Q8SQ23;
DT 01-JUN-2002 (Tremblrel. 21, Created)

DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)

DT 01-JUN-2002 (Tremblrel. 21, last annotation update)
DE T-plasminogen activator.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_Taxid=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ENAMEL ORGAN;
RA Ding Y., Xue J., Bartlett J.D.;
RT "T-plasminogen activator in tooth tissues."
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF364605; AAM00297.1;
SQ SEQUENCE 562 AA; 63668 MW; F9E6B4C77CB101E8 CRC64;

Query Match 42.3%; Score 211; DB 6; Length 562;
Best Local Similarity 48.8%; Pred. No. 1e-17;
Matches 40; Conservative 7; Mismatches 35; Indels 0; Gaps 0;

OY 3 CYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGLGKHNCRPNRR 62
DB 215 CYTGKGLDYRGTRSLTMSGAFCLPWNLSVLGMKITYTAWNSNAOTLGLGKHNCRPNRR 274
OY 63 RPMCXYOVGLKPLVQECMVHDC 84
DB 275 QPMCHVWKDRQLTWECVDPQC 296

RESULT 8

ID Q90675 PRELIMINARY; PRT; 202 AA.
AC Q90675;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, last annotation update)
DE Tissue-type plasminogen activator (Fragment).
GN TPA.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_Taxid=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WHITE LEGHORN;
RX MEDLINE=97199025; PubMed=9047000;
RA Johnson A.L., Bridgman J.T., Anthony R.V.;
RT "Expression of avian urokinase and tissue-type plasminogen activator
messenger ribonucleic acid during follicle development and atresia."
RL Biol. Reprod. 56:581-588(1997).
DR EMBL; U31988; AAA74955.1;
DR HSSP; P00750; 1RTE.
DR MEROPS; S01.232;
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; Kringle_2.
DR Pfam; PF00089; trypsin_1.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringle_2.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 2.
DR PROSITE; PS50240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; UNKNOWN_1.
KW Hydrolyase; Serine protease.
FT NON_TER 1
FT NON_TER 202
SQ SEQUENCE 202 AA; 23148 MW; 049DB42941D83AFC CRC64;

Query Match 38.3%; Score 191; DB 13; Length 202;
Best Local Similarity 43.9%; Pred. No. 9.7e-16;
Matches 36; Conservative 7; Mismatches 39; Indels 0; Gaps 0;

OY 3 CYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGLGKHNCRPNRR 62
DB 41 CYTGKGLDYRGTRSLTMSGAFCLPWNLSVLGMKITYTAWNSNAOTLGLGKHNCRPNRR 100
OY 63 RPMCXYOVGLKPLVQECMVHDC 84
DB 101 QPMCHVWKDRQLTWECVDPQC 122

RESULT 9

ID Q14520 PRELIMINARY; PRT; 560 AA.
AC Q14520; 000663;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, last annotation update)
DE HGF activator like protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kitamura N.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96425001; PubMed=8827452;
RA Choi-Miura N.H., Tobe T., Sumiya J., Nakano Y., Sano Y., Mazda T.,
Tomita M.;
RT "Purification and characterization of a novel hyaluronan-binding
protein (PHBP) from human plasma: it has three EGF, a kringle and a
serine protease domain, similar to hepatocyte growth factor
activator."
RT J. Biochem. 119:1157-1165(1996).
CC -1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
TRYPSIN FAMILY.
DR EMBL; D49742; BAA08576.1;
DR EMBL; S83182; AAB46909.1;
DR HSSP; P00763; IDPO.
DR MEROPS; S01.033;
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF_3.
DR Pfam; PF00051; Kringle_1.
DR Pfam; PF00089; trypsin_1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringle_1.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 3.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolyase; Serine protease.
SQ SEQUENCE 560 AA; 62671 MW; 5C1907230784ACD4 CRC64;

Query Match 37.8%; Score 188.5; DB 4; Length 560;
Best Local Similarity 40.5%; Pred. No. 6.2e-15;
Matches 34; Conservative 16; Mismatches 33; Indels 1; Gaps 1;

OY 3 CYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSNALQLGLGKHNCRPNRR 62
DB 194 CYVGDIYRGKMRKRVNQHACLYWNSHLLQENYMMEDAEHTGIGEHNFCRNPDADE 253
OY 63 RPMCXYOVGLKPLVQECMVHDC 85

Db 254 KPWCFIKYTNDAVKWEYCDVSACS 277

RESULT 10

Q8VCS4 PRELIMINARY; PRT; 653 AA.

AC Q8VCS4; 01-MAR-2002 (TREMBlrel. 20, Created)

DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE Hypothetical 70.6 kDa protein.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_Taxid=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=LIVER;

RA Strausberg R.;

RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC019376; AAH19376.1; -

DR InterPro; IPR001314; Chymotrypsin.

DR InterPro; IPR000561; EGF-like.

DR InterPro; IPR000742; EGF-2.

DR InterPro; IPR001438; EGF-II.

DR InterPro; IPR000083; Fibrinctn1.

DR InterPro; IPR000562; FN_Type-II.

DR InterPro; IPR000001; Kringle.

DR InterPro; IPR001254; Ser_protease_Try.

DR Pfam; PF00008; EGF; 2.

DR Pfam; PF00039; fn1; 1.

DR Pfam; PF00040; fn2; 1.

DR Pfam; PF00051; kringle; 1.

DR Pfam; PF00089; trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR PRINTS; PR00010; EGFBLD.

DR PRINTS; PR00013; FNTYPEIT.

DR PRINTS; PR00018; KRINGLE.

DR ProDom; PD000395; Kringle; 1.

DR ProDom; PD000995; FN_Type-II; 1.

DR SMART; SM00181; EGF; 2.

DR SMART; SM00058; FN1; 1.

DR SMART; SM00059; FN2; 1.

DR SMART; SM00130; KR; 1.

DR SMART; SM00020; Tryp_Spc; 1.

DR PROSITE; PS00022; EGF_1; UNKNOWN_2.

DR PROSITE; PS01186; EGF_2; UNKNOWN_1.

DR PROSITE; PS01253; FIBRONECTIN_1; UNKNOWN_1.

DR PROSITE; PS00023; FIBRONECTIN_2; UNKNOWN_1.

DR PROSITE; PS00021; KRINGLE_1; UNKNOWN_1.

DR PROSITE; PS50070; KRINGLE_2; 1.

DR PROSITE; PS50240; TRYPSIN_DOM; 1.

DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.

DR PROSITE; PS00135; TRYPSIN_SER; UNKNOWN_1.

KW Hypothetical protein.

SQ SEQUENCE 653 AA; 70553 MW; FE18D90174ED6FDD CRC64;

Query Match 37.7%; Score 188; DB 11; Length 653;

Best Local Similarity 53.0%; Pred. No. 8.5e-15;

Matches 35; Conservative 6; Mismatches 25; Indels 0; Gaps 0;

OY 3 CYEGNGHFYRGKASTDTMGRPCLPWN SATVLOQTYHAHRSNALQGLGKHNYCRNPDNR 62

DB 283 CFLNGTEYRGVASTAAGLSCLANSDDLIOELHVDVAAYLLGLGPHAYCRNPDKDE 342

OY 63 RPPWCYV 68

DB 343 RPPWCYV 348

RESULT 11

O35727

ID O35727 PRELIMINARY; PRT; 597 AA.

AC O35727; 01-JAN-1998 (TREMBlrel. 05, Created)

DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE Factor XII.

GN F12.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_Taxid=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=LIVER;

RA Schloesser M., Schwager S., Engel W.;

RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.

DR EMBL; X99571; CAA67891.1; -

DR HSSP; P00760; 1A07.

DR MEROPS; S01.211; -

DR MGD; MGI:1891012; F12.

DR InterPro; IPR001314; Chymotrypsin.

DR InterPro; IPR000561; EGF-like.

DR InterPro; IPR000083; Fibrinctn1.

DR InterPro; IPR000562; FN_Type-II.

DR InterPro; IPR000001; Kringle.

DR InterPro; IPR001254; Ser_protease_Try.

DR Pfam; PF00008; EGF; 2.

DR Pfam; PF00039; fn1; 1.

DR Pfam; PF00040; fn2; 1.

DR Pfam; PF00051; kringle; 1.

DR Pfam; PF00089; trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR PRINTS; PR00018; KRINGLE.

DR ProDom; PD000395; Kringle; 1.

DR ProDom; PD000995; FN_Type-II; 1.

DR SMART; SM00181; EGF; 2.

DR SMART; SM00058; FN1; 1.

DR SMART; SM00059; FN2; 1.

DR SMART; SM00130; KR; 1.

DR SMART; SM00020; Tryp_Spc; 1.

DR PROSITE; PS00022; EGF_1; 2.

DR PROSITE; PS01186; EGF_2; 1.

DR PROSITE; PS01253; FIBRONECTIN_1; 1.

DR PROSITE; PS00023; FIBRONECTIN_2; 1.

DR PROSITE; PS00021; KRINGLE_1; 1.

DR PROSITE; PS50070; KRINGLE_2; 1.

DR PROSITE; PS50240; TRYPSIN_DOM; 1.

DR PROSITE; PS00134; TRYPSIN_HIS; 1.

DR PROSITE; PS00135; TRYPSIN_SER; 1.

KW EGF-like domain; Glycoprotein; Hydrolase; Serine protease.

SQ SEQUENCE 597 AA; 65638 MW; F3AC07C37D0C0FBA CRC64;

Query Match 37.4%; Score 186.5; DB 11; Length 597;

Best Local Similarity 44.0%; Pred. No. 1.2e-14;

Matches 37; Conservative 8; Mismatches 34; Indels 5; Gaps 2;

OY 2 TCYEGNGHFYRGKASTDTMGRPCLPWN SATVLOQTY-HAHRSNALQGLGKHNYCRNPDN 60

DB 216 TCYEGRGSLYRGQAGTQSGAPCCQRW---TYEATYRNMTKEQALSWGIGHAFCRNPDN 271

OY 61 RRRPWCYVQVGLKPLVQECMVHDC 84

DB 272 DTRPWCYVSGDRLSWDYCGLEQC 295

RESULT 12

ID O97507 PRELIMINARY; PRT; 616 AA.

AC O97507; 01-MAY-1999 (TREMBlrel. 10, Created)

DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)

DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)

DT 01-MAR-2002 (Tremblrel. 20, last annotation update)
DE FXII.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Takahashi T., Kihara T.;
RT "Porcine liver factor XII.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
TRYPSIN FAMILY.
DR EMBL; AB022426; BAA37148.1; -.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.211; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000742; EGF-like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR000001; Kring1e.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kring1e; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FNYPEIT.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kring1e; 1.
DR ProDom; PD000995; FN_Type_II; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Serine protease.
SQ SEQUENCE 616 AA; 68012 MW; 4C5FE3D71EBBD1A9 CRC64;

Query Match 35.6%; Score 177.5; DB 6; Length 616;
Best Local Similarity 44.6%; Pred. No. 1.6e-13;
Matches 37; Conservative 6; Mismatches 35; Indels 5; Gaps 2;

OY 3 CYEGNGHYRGKASTDTMGRPLPWN SATVLQOTY-HAHRNALQLGLGKHNYCRNPDNR 61
DB 217 CYSDRGLSYRGMAQTTLGAPCPWAS----EATYWNMTAEQALNMG LGDHAFCRNPDND 272
OY 62 RRPWCYVQVGLKPLVQECMVHDC 84
DB 273 TRPWCFVWRGDQLSWQYCRILARC 295

RESULT 13
ID Q90Y90 PRELIMINARY; PRT; 452 AA.
AC Q90Y90;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, last annotation update)
DE KREMEN.
GN KREMEN.

OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21167372; PubMed=11267660;
RA Nakamura T., Aoki S., Kitajima K., Funakoshi H., Takahashi T.,
RA Matsumoto K., Nakamura T.;
RT "Molecular cloning and characterization of Kremen, a novel kring1e-
RT containing transmembrane protein.";
RL Biochim. Biophys. Acta 1518:63-72(2001).
DR EMBL; AB070851; BAB64294.1; -.
DR InterPro; IPR000859; CUB_domain.
DR InterPro; IPR000001; Kring1e.
DR InterPro; IPR002889; WSC.
DR Pfam; PF00431; CUB; 1.
DR Pfam; PF00051; kring1e; 1.
DR Pfam; PF01822; WSC; 1.
DR ProDom; PD000395; Kring1e; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS00021; KRINGLE_1; UNKNOWN_1.
DR PROSITE; PS50070; KRINGLE_2; 1.
SQ SEQUENCE 452 AA; 50188 MW; ED24BCD1AF4564E2 CRC64;

Query Match 31.5%; Score 157; DB 13; Length 452;
Best Local Similarity 44.1%; Pred. No. 3.9e-11;
Matches 30; Conservative 9; Mismatches 25; Indels 4; Gaps 2;

OY 3 CYEGNGHYRGKASTDTM--GRPLPWN SATVLQOTYHAHRNALQLGLGKHNYCRNPDN 60
DB 30 CYTVNGADYRGQTQNTSLDGKPCLFWNE--TFQHYNTLTKYPNGEGGLGHHNYCRNPDG 87
OY 61 RRPWCYV 68
DB 88 DVSPWCYI 95

RESULT 14
ID Q9ROW3 PRELIMINARY; PRT; 812 AA.
AC Q9ROW3;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, last annotation update)
DE Plasminogen protein precursor (EC 3.4.21.7).
GN PLASMINOGEN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Bangert K., Johnsen A.H., Thorsen S.;
RT "Rat plasminogen: cDNA and gene structure.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=91250378; PubMed=1645711;
RA Kanalas J.J., Makker S.P.;
RT "Identification of the rat Heymann nephritis autoantigen (GP330) as a
RT receptor site for plasminogen.";
RL J. Biol. Chem. 266:10825-10829(1991).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
TRYPSIN FAMILY.
DR EMBL; AJ242649; CAB46014.1; -.
DR HSSP; P00747; IPMK.
DR MEROPS; S01.233; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kring1e.

```

DR      InterPro; IPR0033014; PAN.
DR      InterPro; IPR003609; Pan_app.
DR      InterPro; IPR001254; Ser_protease_Try.
DR      InterPro; IPR001400; Somatotropin.
DR      Pfam; PF00051; kringle; 5.
DR      Pfam; PF00024; PAN; 1.
DR      Pfam; PF00089; trypsin; 1.
DR      PRINTS; PR00722; CHYMOTRYPSIN.
DR      PRINTS; PR00018; KRINGLE.
DR      ProDom; PD000395; kringle; 5.
DR      SMART; SM00130; KR; 4.
DR      SMART; SM00473; PAN_AP; 1.
DR      SMART; SM00020; Tryp_Spc; 1.
DR      PROSITE; PS00021; KRINGLE_1; 5.
DR      PROSITE; PS50070; KRINGLE_2; 5.
DR      PROSITE; PS00338; SOMATOTROPIN_2; UNKNOWN_1.
DR      PROSITE; PS50240; TRYPSIN_DOM; 1.
DR      PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR      PROSITE; PS00135; TRYPSIN_SER; 1.
KW      Hydrolyase; Serine protease; Signal.
FT      SIGNAL      1      19
FT      CHAIN      20      812      PLASMINOGEN.
SQ      SEQUENCE      812 AA; 90535 MW; 8C703C51410EBC9E CRC64;

```

QY 3 CYEGNGHFYRKASTDVTMGPRCLPWNATSATVLOQTYHAHR---SNALQLGLGKHNRYCRND 59

Dd 376 CYOENGKSYRGTSSTNTGTKKCQSW-----VSMTPHSHSKTPANFPDAGL-EMNYCRND 429

QY 60 N-RRRPWCYQVGLKPLY--QECMVHDCAD 86

Dd 430 NDQRGPWCT---TDPSPVRWEYCNLKCSE 456

```

RESULT# 15
O46506
ID O46506 PRELIMINARY; PRT; 454 AA.
AC O46506;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Apolipoprotein a (Fragment).
GN BABAOA.
OS Papio hamadryas (Hamadryas baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Papio.
OX NCBI_TaxID=9557;
RN [1]
RP SEQUENCE FROM N.A.
RA Cox L.A., Jett C., Hixson J.E.;
RT "Molecular Basis of the Apolipoprotein (a) Null Phenotype: A Splice
RT Site Mutation Is Associated with Deletion of a Single Exon in a Null
RT Allele.";
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
CC EMBL; AF029691; AAB97886.1; -.
DR HSSP; P00747; 2PK4.
DR MEROPS; S01.226; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; kringle; 2.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.

```

DR	PROSITE; PS00021; KRINGLE_1; 2.	
DR	PROSITE; PS50070; KRINGLE_2; 2.	
DR	PROSITE; PS50240; TRYPsin_DOM; 1.	
DR	PROSITE; PS00135; TRYPsin_SER; 1.	
KW	Hydrolase; Lipoprotein; Serine protease.	
FT	NON_TER	1
SQ	SEQUENCE	454 AA; 50041 MW; 974E30744C187B2F CRC64;

Query Match	30.7%;	Score 153;	DB 6;	Length 454;
Best Local Similarity	38.6%;	Pred. No. 1.2e-10;		
Matches 34;	Conservative 11;	Mismatches 31;	Indels 12;	Gaps 4;

QY 3 CYEGNGHFRTGKASTDITMGRPCLPMNSATVLQ--QTYNHAKRSNALQIGLGNHYCRNPND 60
 ||| : ||| | | : | : | : |||||
Db 102 CYHGDSYRGSEFTVTGTGTCQSMSSMTPHQHKRTRENPRNDGLTM-----NYCRNPDA 156

```

QY      61 RRRPWCYVQVGLKPLY--QECMVHDCAD 86
          |||  :| | :| :| |
Db     157 DTGPMCF--MDPSVRWEYCNLTRCSD 181

```

Search completed: November 11, 2002, 13:01:56
Job time : 78 secs

•
•
•
•

• •

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 11, 2002, 11:35:58 ; Search time 2622 Seconds
(without alignments)
2863.663 Million cell updates/sec

Title: PCT-US02-27855-2
Perfect score: 258
Sequence: 1 aaaacctgctatgaggggaa.....tggtgcatgactgcagat 258

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_un:*
28: em_vl:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. 1s the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	256.4	99.4	423	6 A76865	A76865 Sequence 3
2	256.4	99.4	1236	6 E02578	E02578 DNA encodin
3	256.4	99.4	1236	6 E02579	E02579 Human prout
4	256.4	99.4	1236	6 E02708	E02708 DNA sequenc
5	256.4	99.4	1236	6 E02709	E02709 DNA sequenc
6	256.4	99.4	1236	6 E02710	E02710 DNA sequenc
7	256.4	99.4	1236	6 E02832	E02832 DNA encodin
8	256.4	99.4	1236	6 E02833	E02833 DNA encodin
9	256.4	99.4	1296	6 E02577	E02577 DNA encodin
10	256.4	99.4	1296	6 E02647	E02647 DNA sequenc
11	256.4	99.4	1296	6 E02649	E02649 DNA sequenc
12	256.4	99.4	1296	6 E02711	E02711 DNA sequenc
13	256.4	99.4	1296	6 E06064	E06064 DNA encodin
14	256.4	99.4	1359	6 AX451990	AX451990 Sequence
15	256.4	99.4	1394	6 E00421	E00421 CDNA coding
16	256.4	99.4	1473	6 E02114	E02114 CDNA sequen
17	256.4	99.4	1474	6 E00924	E00924 CDNA encodi
18	256.4	99.4	1474	6 E01559	E01559 CDNA sequen
19	256.4	99.4	1474	6 E01580	E01580 CDNA encodi
20	256.4	99.4	1474	6 E02095	E02095 DNA sequenc
21	256.4	99.4	1475	6 AX365729	AX365729 Sequence
22	256.4	99.4	1475	6 E01560	E01560 CDNA sequen
23	256.4	99.4	1475	6 HUMUKM1	K03226 Human prepr
24	256.4	99.4	1964	6 A18397	A18397 Human uPA c
25	256.4	99.4	1964	6 AX402384	AX402384 Sequence
26	256.4	99.4	1964	6 AX451989	AX451989 Sequence
27	256.4	99.4	2293	6 A11978	A11978 Synthetic n
28	256.4	99.4	2294	6 A09202	A09202 Artificial
29	256.4	99.4	2294	6 AX365730	AX365730 Sequence
30	256.4	99.4	2294	9 HUMUKA	D00244 Homo sapien
31	256.4	99.4	2294	11 G27040	G27040 SHGC-31374
32	256.4	99.4	2296	6 A35395	A35395 H.sapiens u
33	256.4	99.4	2300	6 A04029	A04029 Synthetic H
34	256.4	99.4	2300	6 A21571	A21571 pro-Urokina
35	256.4	99.4	2301	6 E01467	E01467 DNA encodin
36	256.4	99.4	2303	6 I07013	I07013 Sequence 4
37	256.4	99.4	2303	6 I08092	I08092 Sequence 1
38	256.4	99.4	2304	6 E00178	E00178 CDNA encodi
39	256.4	99.4	2304	6 I03932	I03932 Sequence 6
40	256.4	99.4	2304	6 I04632	I04632 Sequence 1
41	256.4	99.4	2304	9 HUMUKPM	M15476 Human pro-u
42	256.4	99.4	2345	9 BC013575	BC013575 Homo sapi
43	256.4	99.4	2629	9 HUMUK	K02286 Human urok1
44	256.4	99.4	10332	6 A83180	A83180 Sequence 13
45	254.8	98.8	473	9 AY029537	AY029537 Homo sapi

ALIGNMENTS

RESULT 1
A76865 LOCUS A76865 423 bp DNA linear PAT 19-OCT-1999
DEFINITION Sequence 3 from Patent WO9315199.
ACCESSION A76865
VERSION A76865.1 GI:6088666
KEYWORDS
SOURCE unidentified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 423)
AUTHORS Fleer, R. and Fournier, A.
TITLE NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES, PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION CONTAINING SAID POLYPEPTIDES
JOURNAL Patent: WO 9315199-A 3 05-AUG-1993;

FEATURES	RHONE	POULENC	RORER	SA (FR)
source	1..423	Location/Qualifiers		
CDS		/organism="unidentified"		
		/db_xref="taxon:32644"		
		<3..419		
		/codon_start=1		
		/product="Fragment C-ter de la chimere SAH-UK1-135"		
		/protein_id="CAB58681.1"		
		/db_xref="GI:6088667"		
		/translation="IGLSNELHQVPSNCDLNGTCVSNKYFYSNIHMCNPKKFGGQH CEIDSKTCYEGNGHYRGKASTDIDMRCLPWNASTATVLOQTYHAHRSDALQLGLKH NYCRNPDNRHRPWCYVQVGLKPLVQECMVHDCADGK"		
BASE COUNT	117 a	109 c	107 g	90 t
ORIGIN				
Query Match	99.4%	Score 256.4;	DB 6;	Length 423;
Best Local Similarity	99.6%	Pred. No. 1.7e-66;		
Matches 257;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY 1	AAACCTGCTATGAGGGGATGTGTCACCTTTACCGAGGAAGGCCAGCAGTACACCATG 60			
DB 153	AAAACCTGCTATGAGGGGATGTGTCACCTTTACCGAGGAAGGCCAGCAGTACACCATG 212			
QY 61	GGCCGGCCCTGCGCTGCGCCCTGGGAAGTCTGCGCACTGTCTTGACGAAAGTACCATGCCCCAC 120			
DB 213	GGCCGGCCCTGCGCTGCGCCCTGGGAAGTCTGCGCACTGTCTTGACGAAAGTACCATGCCCCAC 272			
QY 121	AGATCTATGCTCTTTCAGCTGGGCGCTGGGGAACATAATTAAGTCAAGAACCCAGACCAAC 180			
DB 273	AGATCTATGCTCTTTCAGCTGGGCGCTGGGGAACATAATTAAGTCAAGAACCCAGACCAAC 332			
QY 181	CGGAGGCGGACCCCTGCTGCTATGTGCAAGGTGGGCGCTAAAGCCGCTTGTCCAAGAGTGCATG 240			
DB 333	CGGAGGCGGACCCCTGCTGCTATGTGCAAGGTGGGCGCTAAAGCCGCTTGTCCAAGAGTGCATG 392			
QY 241	GTCATGACTGCGCAGAT 258			
DB 393	GTCATGACTGCGCAGAT 410			
RESULT 2				
E02578	LOCUS	1236 bp	DNA	linear
E02578	DEFINITION	DNA encoding human prourokinase derivative.		PAT 29-SEP-1997
E02578	ACCESSION	E02578		
E02578.1	VERSION	GI:2170808		
JP 1990227075-A/6.	KEYWORDS			
JP 1990227075-A/6.	SOURCE			
synthetic construct.	ORGANISM			
synthetic construct.	REFERENCE			
artificial sequences.	1 (bases 1 to 1236)			
	Sasaki,K., Nishi,T., Yasumura,S., Sato,M. and Itou,S.			
	NEW POLYPEPTIDE			
	Patent: JP 1990227075-A 6 10-SEP-1990;			
	KYOMA HAKKO KOGYO CO LTD			
	OS Artificial gene			
	OC Artificial sequence; Genes.			
	PN JP 1990227075-A/6			
	PD 10-SEP-1990			
	PF 28-SEP-1989 JP 1989253097			
	PR 29-SEP-1988 JP 88P 245705			
	PI SASAKI KATSUTOSHI, NISHI TATSUYA, YASUMURA SHIGEYOSHI, PI			
	SATO MORIYUKI,			
	PI ITOU SEIGA			
	PC C12N9/72, C07K13/00, C07K15/14, C12N1/21, C12N5/10, C12N9/64, PC			
	C12N15/27,			
	PC C12N15/58, C12N15/70, C12N15/85, C12P21/02, C12P21/02, (C12N1/21,			
	PC C12R1:19),			
	PC (C12P21/02, C12R1:19), (C12P21/02, C12R1:91);			
	CC strandedness: Double;			
	CC topology: Linear;			
	CC hypothetical: No;			

CC	anti-sense: No;	Location/Qualifiers	
FH	key		
FH			
FT	mat_peptide	1..1233	
	derivative(proVK-S1) which is FT		resistant to
FEATURES	protease'	Location/Qualifiers	
source	1..1236		
	/organism="synthetic construct"		
	/db_xref="taxon:32630"		
BASE COUNT	321 a 341 c 318 g 256 t		
ORIGIN			
Query Match	99.4%;	Score 256.4;	DB 6;
Best Local Similarity	99.6%;	Pred. No. 1.6e-66;	Length 1236;
Matches	257;	Conservative	0; Mismatches 1; Indels 0; Gaps 0;
QY	1	AAACCTGCTATGAGGGGAATGGTCACTTTTACCAGAGAAAGGCCAGCAGTGCACCATG	60
Db	142	AAACCTGCTATGAGGGGAATGGTCACTTTTACCAGAGAAAGGCCAGCAGTGCACCATG	201
QY	61	GCGCGGCCCTGCCTGCCTCGGAAGTCTGCGCACTGTCTTCAGCAAAAGTACCATGCCCCAC	120
Db	202	GCGCGGCCCTGCCTGCCTCGGAAGTCTGCGCACTGTCTTCAGCAAAAGTACCATGCCCCAC	261
QY	121	AGATCTAATGCTCTTTCAGCTGGGCGCTGGGGAACATAATTAATGTCAGAGAACCCAGACAAC	180
Db	262	AGATCTGATGCTCTTTCAGCTGGGCGCTGGGGAACATAATTAATGTCAGAGAACCCAGACAAC	321
QY	181	CGGAGCGGACCCCTGGTGCCTATGTGCAGAGTGGGGCTAAAGCCGCTTGTCCAAAGAGTGCATG	240
Db	322	CGGAGCGGACCCCTGGTGCCTATGTGCAGAGTGGGGCTAAAGCCGCTTGTCCAAAGAGTGCATG	381
QY	241	GTGCATGACTGCGCAGAT	258
Db	382	GTGCATGACTGCGCAGAT	399
RESULT 3			
E02579		1236 bp	DNA
LOCUS	E02579		linear
DEFINITION	Human prourokinase derivative.		PAT 29-SEP-1997
ACCESSION	E02579		
VERSION	E02579.1	GI:2170809	
KEYWORDS	JP 1990227075-A/7.		
SOURCE	synthetic construct.		
ORGANISM	synthetic construct.		
REFERENCE	1 (bases 1 to 1236)		
AUTHORS	Sasaki,K., Nishi,T., Yasumura,S., Sato,M. and Itou,S.		
TITLE	NEW POLYPEPTIDE		
JOURNAL	Patent: JP 1990227075-A 7 10-SEP-1990;		
COMMENT	KYOWA HAKKO KOGYO CO LTD		
	OS Artificial gene		
	OC Artificial sequence; Genes.		
	PN JP 1990227075-A/7		
	PD 10-SEP-1990		
	PF 28-SEP-1989 JP 1989253097		
	PR 29-SEP-1988 JP 88P 245705		
	PI SASAKI KATSUTOSHI, NISHI TATSUYA, YASUMURA SHIGEYOSHI, PI		
	SATO MORIYUKI,		
	PI ITOU SEIGA		
	PC C12N9/72,C07K13/00,C07K15/14,C12N1/21,C12N5/10,C12N9/64, PC		
	C12N15/27,		
	PC C12N15/58,C12N15/70,C12N15/85,C12P21/02,C12P21/02,(C12N1/21,		
	PC C12R1:19),		
	PC (C12P21/02,C12R1:19),(C12P21/02,C12R1:91);		
	CC strandedness: Double;		
	CC topology: Linear;		
	CC hypothetical: No;		
	CC anti-sense: No;		
	FH key	Location/Qualifiers	

Db	142	AAACCTGCTATGAGGGGAATG	GTCACTTTTACCAGGAAGGCCAGCACTGACACCATG	201	
QY	61	GGCCGGCCCTG	CCCTGGAACTCTGCCACTGTCTT	CAGCAAAACGTACCATG	CCAC
Db	202	GGCCGGCCCTG	CCCTGGAACTCTGCCACTGTCTT	CAGCAAAACGTACCATG	CCAC
QY	121	AGATCTAATGCTTTCAGCTGGGC	CTGGGAAACATATATTACTG	CAGGAACCCAGACAC	180
Db	262	AGATCTGATGCTTTCAGCTGGGC	CTGGGAAACATATATTACTG	CAGGAACCCAGACAC	321
QY	181	CGGAGGCGACCCCTGTGCTATGTG	CAGGTGGGCTTAAAGCCGCTGTG	CCAAGAGTGCATG	240
Db	322	CGGAGGCGACCCCTGTGCTATGTG	CAGGTGGGCTTAAAGCCGCTGTG	CCAAGAGTGCATG	381
QY	241	GTGCATGACTGCCAGAT	258		
Db	382	GTGCATGACTGCCAGAT	399		

RESULT 6

LOCUS	E02710	1236 bp	DNA	linear	PAT 29-SEP-1997
DEFINITION	DNA sequence coding for prourokinase derivative, UK-S3.				
ACCESSION	E02710				
VERSION	E02710.1	GI:2170938			
KEYWORDS	JP 1991022979-A/3.				
SOURCE	Unidentified.				

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	(bases 1 to 1236)	Yasumura, S., Nishi, T. and Itou, S.	NOVEL PLASMINOGEN-ACTIVATION FACTOR	
		Patent: JP 1991022979-A 3 31-JAN-1991;		
		KYOWA HAKKO KOGYO CO LTD		
		PN JP 1991022979-A/3		

COMMENT

```

PF 19-JUN-1989 JP 1989J156302
PI YASUMURA SHIGEYOSHI, NISHI TATSUYA, ITOU SEIGA PC
C12N9/72,C12N15/58,(C12N9/72,C12R1:19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: library=plasmid PUKS3;
FH Key Location/Qualifiers
FT CDS 1..1236
FT modified_base 457..459 /product='Prourokinase derivative, UK-S3' FT
FT modified_base 463..465 /note='ctg' is replaced by 'aat' FT
FT /note='ccc' is replaced by 'acg'
FEATURES
source 1..1236

```

nos

BASE COUNT	322	a	338	c	317	g	259	t
ORIGIN	/organism="Hepatitis B virus" /db_xref="taxon:10407"							

Query Match	99.4%;	Score 256.4;	DB 6;	Length 1236;
Best Local Similarity	99.6%;	Pred. No. 1.6e-66;		
Matches 257;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;

[illegible]

Db	262	AGATCTGATGCTCTTCACGCTGGGCGCTGGGGAACATATTAATTACTGCAGGAACCCAGACAAC	321
QY	181	CGAGGCGACCTGTGTCTATGTGCAGTGGGGCTAAAGCCGCTGTGCCAAGATGCATG	240
Db	322	CGGAGCGACCCCTGGTGTCTATGTGCAGTGGGGCTAAAGCCGCTGTGCCAAGATGCATG	381
QY	241	GTGCATGACTGGCGCAGAT	258
Db	382	GTGCATGACTGGCGCAGAT	399

RESULT 7

LOCUS	E02832	1236 bp	RNA	linear	PAT 29-SEP-1997
DEFINITION	DNA encoding human pro-urokinase.				
ACCESSION	E02832				
VERSION	E02832.1	GI:2171060			
KEYWORDS	JP 1991087180-A/1.				
SOURCE	Homo sapiens.				
ORGANISM	Homo sapiens				

REFERENCE	AUTHORS	TITLE	JOURNAL
1 (bases 1 to 1236)	Tanabe, T., Amatsuji, Y., Kasai, S., Hirose, M., Morita, M., Kawabe, H. and Arimura, H.	MUTANT HUMAN-PROTEUKINASE, ITS PRODUCTION, DNA SEQUENCE, PLASMID AND HOST	Patent: JP 1991087180-A 1 11-APR-1991;

JOURNAL

COMMENT	OS	Homo sapiens (human)
	PN	JP 1991087180-A/1
	PD	11-APR-1991
	PF	16-FEB-1990 JP 1990036809
	PR	18-MAY-1989 JP 89P 126434
	PI	TANABE TOSHIZUMI, AMATSUJI YASUO, KASAI SHUNJI, HIROSE MASAKI, PI MORITA MASANORI, KAWABE HARUHIIDE, ARIMURA HIROBUMI PC
		C12N9/72, C12N5/10, C12N15/58;

FEATURES	source
CC	topology: Linear;
CC	hypothetical: No;
CC	anti-sense: No;
FH	Key
FT	CDS
FT	mat_peptide
FT	mutation
FT	replace(70..72,'AAC')
FT	replace(85..87,'TCC')
FT	1..147
FT	location/Qualifiers
FT	1..1236

nos

```
BASE COUNT      318 a      342 c      319 g      257 t
ORIGIN
/organism="Homo sapiens"
/db_xref="taxon:9606"
```

Query Match	99.48;	Score 256.4;	DB 6;	Length 1236;
Best Local Similarity	99.68;	Pred. NO. 1.6e-66;		
Matches 257;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;

[illegible]

```
QY 121 AGATCTAATGCTCTTCAGCTGGGCGTGGGGAACATAATTACTGCAGAACCCAGACAAC 180
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 262 AGATCTGATGCTCTTCAGCTGGGCGTGGGGAACATAATTACTGCAGAACCCAGACAAC 321
QY 181 CGAGGCGACCCCTGTGCTATGTGCAGAGTGGGCTAAAGCCGCTGTGCCAAGAGTGCATG 240
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 322 CGAGGCGACCCCTGTGCTATGTGCAGAGTGGGCTAAAGCCGCTGTGCCAAGAGTGCATG 381
QY 241 GTGCATGACTGCCGAGAT 258
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 382 GTGCATGACTGCCGAGAT 399

RESULT 8
E02833 1236 bp RNA linear PAT 29-SEP-1997
LOCUS DNA encoding human pro-urokinase.
DEFINITION E02833
ACCESSION E02833
VERSION E02833.1 GI:2171061
KEYWORDS JP 1991087181-A/1.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1236)
AUTHORS Tanabe,T., Amatsuji,Y., Kasai,S., Hirose,M., Morita,M., Kawabe,H.
          and Arimura,H.
TITLE MUTANT HUMAN-PROUROKINASE, ITS PRODUCTION, DNA SEQUENCE, PLASMID
          AND HOST
JOURNAL Patent: JP 1991087181-A 1 11-APR-1991;
          GREEN CROSS CORP.:THE
COMMENT OS Homo sapiens (human)
          PN JP 1991087181-A/1
          PD 11-APR-1991
          PF 22-FEB-1990 JP 1990042020
          PR 18-MAY-1989 JP 89P 126433
          PI TANABE TOSHIZUMI, AMATSUJI YASUO, KASAI SHUNJI, HIROSE
          MASAOKI, PI MORITA MASANORI, KAWABE HARUHIDE, ARIMURA HIROBUMI PC
          C12N9/72,C12N5/10,C12N15/58;
          CC strandedness: Double;
          CC topology: Linear;
          CC hypothetical: No;
          CC anti-sense: No;
          FH Key Location/Qualifiers
          FH CDS 1..1236
          FT /product='pro-urokinase'
          FT mat_peptide 1..1233
          FT /product='pro-urokinase'
          FT mutation replace(31..96,'')
          FT /product='pro-urokinase mutant' FT mutation
          FT replace(97..126,'')
          FT /product='pro-urokinase mutant' FT Region
          FT 1..147 /note='epidermal growth factor domain'.
FEATURES
source 1..1236
          /organism="Homo sapiens"
          /db_xref="taxon:9606"
BASE COUNT 318 a 342 c 319 g 257 t
ORIGIN
Query Match 99.4%; Score 256.4; DB 6; Length 1236;
Best Local Similarity 99.6%; Pred. No. 1.6e-66;
Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Db 202 GGCCGCCCTGCTGCTGCCCTGGAACTCTGCCACTGTCTTCAGCAAAACGTACCATGCCAC 261
QY 121 AGATCTAATGCTCTTCAGCTGGGCGTGGGGAACATAATTACTGCAGAACCCAGACAAC 180
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 262 AGATCTGATGCTCTTCAGCTGGGCGTGGGGAACATAATTACTGCAGAACCCAGACAAC 321
QY 181 CGAGGCGACCCCTGTGCTATGTGCAGAGTGGGCTAAAGCCGCTGTGCCAAGAGTGCATG 240
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 322 CGAGGCGACCCCTGTGCTATGTGCAGAGTGGGCTAAAGCCGCTGTGCCAAGAGTGCATG 381
QY 241 GTGCATGACTGCCGAGAT 258
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 382 GTGCATGACTGCCGAGAT 399

RESULT 9
E02577 1296 bp RNA linear PAT 29-SEP-1997
LOCUS DNA encoding human prourokinase.
DEFINITION E02577
ACCESSION E02577
VERSION E02577.1 GI:2170807
KEYWORDS JP 1990227075-A/5.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1296)
AUTHORS Sasaki,K., Nishi,T., Yasumura,S., Sato,M. and Itou,S.
          NEW POLYPEPTIDE
TITLE Patent: JP 1990227075-A 5 10-SEP-1990;
          KYOWA HAKKO KOGYO CO LTD
JOURNAL OS Homo sapiens (human)
          PN JP 1990227075-A/5
          PD 10-SEP-1990
          PF 28-SEP-1989 JP 1989253097
          PR 29-SEP-1988 JP 88P 245705
          PI SASAKI KATSUTOSHI, NISHI TATSUYA, YASUMURA SHIGEYOSHI, PI
          SATO MORIYUKI,
          PI ITOU SEIGA
          PC C12N9/72,C07K13/00,C07K15/14,C12N1/21,C12N5/10,C12N9/64, PC
          C12N15/27,
          PC C12N15/58,C12N15/70,C12N15/85,C12P21/02,C12P21/02,(C12N1/21,
          PC C12R1:19),
          PC (C12P21/02,C12R1:19),(C12P21/02,C12R1:91);
          CC strandedness: Double;
          CC topology: Linear;
          CC hypothetical: No;
          CC anti-sense: No;
          CC *source: cell_type=Detroit 562 cell;
          CC *source: clone=DVK11;
          FH Key Location/Qualifiers
          FH CDS 1..1296
          FT /product='human prourokinase precursor' FT
          FT mat_peptide 61..1293
          FT /product='human peourokinase'.
FEATURES
source 1..1296
          /organism="Homo sapiens"
          /db_xref="taxon:9606"
BASE COUNT 327 a 361 c 337 g 271 t
ORIGIN
Query Match 99.4%; Score 256.4; DB 6; Length 1296;
Best Local Similarity 99.6%; Pred. No. 1.6e-66;
Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```


Db 262 GGGCGGCGCTGCTGCCCCCTGGAACCTCTGCCACTGTCTTCAGCAAAAGCTACCATGCCAC 321

QY 121 AGATCTAATGCTCTTCAGCTGGGCGCTGGGAAACATATATCTGAGGAACCCAGACAAC 180

Db 322 AGATCTGATGCTCTTCAGCTGGGCGCTGGGAAACATATATCTGAGGAACCCAGACAAC 381

QY 181 CGGAGCGGACCCCTGCTATGTGCAAGTGGGCTAAAGCCGCTTGTCCAAAGAGTGCATG 240

Db 382 CGGAGCGGACCCCTGCTATGTGCAAGTGGGCTAAAGCCGCTTGTCCAAAGAGTGCATG 441

QY 241 GTGCATGACTGCGCAGAT 258

Db 442 GTGCATGACTGCGCAGAT 459

RESULT 10

LOCUS E02647 1296 bp RNA linear PAT 29-SEP-1997

DEFINITION DNA sequence coding for pro-urokinase.

ACCESSION E02647

VERSION E02647.1 GI:2170875

KEYWORDS JP 1990257891-A/2.

SOURCE Homo sapiens.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 1296)

AUTHORS Miyajima, H., Sasaki, K. and Ito, S.

TITLE PRODUCTION OF PROTEIN BY RECOMBINANT ANIMAL CELL

JOURNAL Patent: JP 1990257891-A 2 18-OCT-1990; KYOWA HAKKO KOGYO CO LTD OS Homo sapiens (human) PN JP 1990257891-A/2 PD 18-OCT-1990 PF 31-MAR-1989 JP 1989078573 PI MIYAJI HIROMASA, SASAKI KATSUNOSHI, ITOU SEIGA PC C12P21/00, C12P21/02//C12N5/10, C12N15/09, (C12P21/00, C12R1:91), PC (C12N5/10, C12R1:91);

FEATURES

source location/Qualifiers

1..1296 /product='pro-urokinase'.

BASE COUNT 327 a 361 c 337 g 271 t

ORIGIN

Query Match 99.4%; Score 256.4; DB 6; Length 1296; Best Local Similarity 99.6%; Pred. No. 1.6e-66; Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAACCTGCTATGAGGGGAATGCTCACTTTTACCGAGGAAGGCCAGACTGACACCATG 60

Db 202 AAAACCTGCTATGAGGGGAATGCTCACTTTTACCGAGGAAGGCCAGACTGACACCATG 261

QY 61 GGGCGGCGCTGCTGCCCCCTGGAACCTCTGCCACTGTCTTCAGCAAAAGCTACCATGCCAC 120

Db 262 GGGCGGCGCTGCTGCCCCCTGGAACCTCTGCCACTGTCTTCAGCAAAAGCTACCATGCCAC 321

QY 121 AGATCTAATGCTCTTCAGCTGGGCGCTGGGAAACATATATCTGAGGAACCCAGACAAC 180

Db 322 AGATCTGATGCTCTTCAGCTGGGCGCTGGGAAACATATATCTGAGGAACCCAGACAAC 381

QY 181 CGGAGCGGACCCCTGCTATGTGCAAGTGGGCTAAAGCCGCTTGTCCAAAGAGTGCATG 240

Db 382 CGGAGCGGACCCCTGCTATGTGCAAGTGGGCTAAAGCCGCTTGTCCAAAGAGTGCATG 441

QY 241 GTGCATGACTGCGCAGAT 258

Db 442 GTGCATGACTGCGCAGAT 459

RESULT 11

LOCUS E02649 1296 bp DNA linear PAT 29-SEP-1997

DEFINITION DNA sequence coding for human pro-urokinase.

ACCESSION E02649

VERSION E02649.1 GI:2170877

KEYWORDS JP 1990261386-A/1.

SOURCE Homo sapiens.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 1296)

AUTHORS Sekine, S., Ito, S. and Katsuki, M.

TITLE RECOMBINANT VECTOR

JOURNAL Patent: JP 1990261386-A 1 24-OCT-1990; KYOWA HAKKO KOGYO CO LTD, JITSUKEN DOUBUTSU CHUO KENKYUSHO OS Homo sapiens (human) PN JP 1990261386-A/1 PD 24-OCT-1990 PF 31-MAR-1989 JP 1989078574 PI SEKINE SUSUMU, ITOU SEIGA, KATSUKI MOTOYA PC C12N15/85, A01K67/027, C12N1/21, C12N9/72//C12N15/58, (C12N1/21, C12R1:19);

FEATURES

source location/Qualifiers

1..1296 /product='Human pro-urokinase'.

BASE COUNT 327 a 360 c 338 g 271 t

ORIGIN

Query Match 99.4%; Score 256.4; DB 6; Length 1296; Best Local Similarity 99.6%; Pred. No. 1.6e-66; Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAACCTGCTATGAGGGGAATGCTCACTTTTACCGAGGAAGGCCAGACTGACACCATG 60

Db 202 AAAACCTGCTATGAGGGGAATGCTCACTTTTACCGAGGAAGGCCAGACTGACACCATG 261

QY 61 GGGCGGCGCTGCTGCCCCCTGGAACCTCTGCCACTGTCTTCAGCAAAAGCTACCATGCCAC 120

Db 262 GGGCGGCGCTGCTGCCCCCTGGAACCTCTGCCACTGTCTTCAGCAAAAGCTACCATGCCAC 321

QY 121 AGATCTAATGCTCTTCAGCTGGGCGCTGGGAAACATATATCTGAGGAACCCAGACAAC 180

Db 322 AGATCTGATGCTCTTCAGCTGGGCGCTGGGAAACATATATCTGAGGAACCCAGACAAC 381

QY 181 CGGAGCGGACCCCTGCTATGTGCAAGTGGGCTAAAGCCGCTTGTCCAAAGAGTGCATG 240

Db 382 CGGAGCGGACCCCTGCTATGTGCAAGTGGGCTAAAGCCGCTTGTCCAAAGAGTGCATG 441

QY 241 GTGCATGACTGCGCAGAT 258

Db 442 GTGCATGACTGCGCAGAT 459

RESULT 12

LOCUS E02711 1296 bp DNA linear PAT 29-SEP-1997

DEFINITION	DNA sequence coding for native prourokinase.				
ACCESSION	E02711				
VERSION	E02711.1 GI:2170939				
KEYWORDS	JP 1991022979-A/4.				
SOURCE	unidentified.				
ORGANISM	Hepatitis B virus				
REFERENCE	1 (bases 1 to 1296)				
AUTHORS	Yasumura,S., Nishi,T. and Ito,S.				
TITLE	NOVEL PLASMINOGEN-ACTIVATION FACTOR				
JOURNAL	Patent: JP 1991022979-A 4 31-JAN-1991;				
	KYOWA HAKKO KOGYO CO LTD				
COMMENT	PN JP 1991022979-A/4				
	PD 31-JAN-1991				
	PF 19-JUN-1989 JP 1989156302				
	PI YASUMURA SHIGEYOSHI, NISHI TATSUYA, ITOU SEIGA PC				
	C12N9/72,C12N15/58,(C12N9/72,C12R1.19);				
	CC strandedness: Double;				
	CC topology: Linear;				
	CC hypothetical: No;				
	CC anti-sense: No;				
	FH Key Location/Qualifiers				
	FT	sig_peptide	1..60	/product='Native prourokinase signal peptide'	
	FT		61..1296		
	FT	CDS	/product='Native prourokinase'	FT	
	FT		mat_peptide	61..192	
	FT			/note='Growth factor domain'	
	FT		mat_peptide	193..459	
	FT			/note='taringul domain'	
	FT		mat_peptide	460..1293	
	FT			/note='Protease domain'.	
FEATURES	Location/Qualifiers				
SOURCE	1..1296				
	/organism='Hepatitis B virus'				
	/db_xref='taxon:10407'				
BASE COUNT	327 a	361 c	337 g	271 t	
ORIGIN					
Query Match	99.4%;	Score 256.4;	DB 6;	Length 1296;	
Best Local Similarity	99.6%;	Pred. No. 1.6e-66;			
Matches 257;	Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	1	AAACCTGCTATGAGGGGAATGCTCACTTTTACCGAGGAAGCCAGCACTGACACCATG	60		
Db	202	AAACCTGCTATGAGGGGAATGCTCACTTTTACCGAGGAAGCCAGCACTGACACCATG	261		
QY	61	GCGCGGCCCTGCGCTGCCCTGGAATCTGCCACTGTCTTCAGCAAAACGTACCATGCCAC	120		
Db	262	GCGCGGCCCTGCGCTGCCCTGGAATCTGCCACTGTCTTCAGCAAAACGTACCATGCCAC	321		
QY	121	AGATCTAATGCTCTTCAGCTGGGCCCTGGGGAACATAATATTACTGCAGGAACCCAGACAAC	180		
Db	322	AGATCTGATGCTCTTCAGCTGGGCCCTGGGGAACATAATATTACTGCAGGAACCCAGACAAC	381		
QY	181	CGGAGCGCAACCCCTGGTGCCTATGTGCAGAGGTGGGCCCTAAAGCCGCTTGTCCTCAAGAGTGCA	240		
Db	382	CGGAGCGCAACCCCTGGTGCCTATGTGCAGAGGTGGGCCCTAAAGCCGCTTGTCCTCAAGAGTGCA	441		
QY	241	GTGCATGACTGCGCAGAT	258		
Db	442	GTGCATGACTGCGCAGAT	459		
RESULT	13				
LOCUS	E06064	1296 bp	RNA	linear	PAT 29-SEP-1997
DEFINITION	DNA encoding human urokinase.				
ACCESSION	E06064				
VERSION	E06064.1 GI:2174251				
KEYWORDS	JP 1993336965-A/2.				
SOURCE	Homo sapiens.				

ORGANISM	Homo sapiens				
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.				
AUTHORS	Yasumura, S., Yamamoto, Y., Hasegawa, M., Higo, K., Kubo, K. and Kuwabara, T.				
TITLE	HUMAN PROUROKINASE DERIVATIVE				
JOURNAL	Patent: JP 199336965-A 2 21-DEC-1993;				
COMMENT	KYOWA HAKKO KOGYO CO LTD OS Homo sapiens (human) PN JP 199336965-A/2 PD 21-DEC-1993 PF 17-OCT-1991 JP 1991269615 PI YASUMURA SHIGEYOSHI, YAMAMOTO YOSHINORI, HASEGAWA MAMORU, PI HIGO KATSUYA, PI KUBO KAZUHIRO, KUWABARA TAKASHI PC C12N9/72, C12N1/21, C12N15/58; CC strandedness: Double; CC topology: Linear; CC *source: strain=Detroit562; CC *source: clone=PUK1, PUK11; FH Key FH Location/Qualifiers FT CDS FT 1. .1296 FT /product='human prourokinase' FT sig_peptide FT 1. .60 FT mutation FT /replace(88. .195, ''') FT /note='UK-deltaGSI' FT mutation FT /replace(540. .541, 'aa') FT /note='UK-deltaGSI'. FT Location/Qualifiers FT 1. .1296 FT /organism='Homo sapiens' FT /db_xref='taxon:9606'				
FEATURES	source				
BASE COUNT	327 a 361 c 337 g 271 t				
ORIGIN					
Query Match	99.4%; Score 256.4; DB 6; Length 1296;				
Best Local Similarity	99.6%; Pred. No. 1.6e-66;				
Matches	257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;				
QY	1	AAACCTGCTATGAGGGGAATGCTACTTTTACCGAGAAAGGCCAGACTGACACCATG	60		
Db	202	AAACCTGCTATGAGGGGAATGCTACTTTTACCGAGAAAGGCCAGACTGACACCATG	261		
QY	61	GGCCGGCCCTGCGCTGCGCTGGAACCTCTGCCACTGTCCCTTCAGCAAAACGTACCATGCCAC	120		
Db	262	GGCCGGCCCTGCGCTGCGCTGGAACCTCTGCCACTGTCCCTTCAGCAAAACGTACCATGCCAC	321		
QY	121	AGATCTAATGCTCTTCAGCTGCGCCCTGGGGAACATTAATTACTGCAGGAACCCAGACAAC	180		
Db	322	AGATCTAATGCTCTTCAGCTGCGCCCTGGGGAACATTAATTACTGCAGGAACCCAGACAAC	381		
QY	181	CGAGAGCGACCCCTGCTGCTATGTCAGGTGGGCTAAAGCCCTGTCCAAAGAGTGCATG	240		
Db	382	CGAGAGCGACCCCTGCTGCTATGTCAGGTGGGCTAAAGCCCTGTCCAAAGAGTGCATG	441		
QY	241	GTGCATGACTGCGCAGAT	258		
Db	442	GTGCATGACTGCGCAGAT	459		
RESULT 14					
LOCUS	AX451990	1359 bp	DNA	linear	PAT 06-JUL-2002
DEFINITION	Sequence 3 from Patent WO0244393.				
ACCESSION	AX451990				
VERSION	AX451990.1	GI:21711991			
KEYWORDS	human.				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.				

Tue Nov 12 18:06:49 2002

pct-us02-27855-2.rge

Page 8

REFERENCE	1	Armenclariiz Borunda,J. and Aguilar Cordova,E.
AUTHORS		Recombinant viral and non-viral vectors containing the human
TITLE		urokinase plasminogen activator gene and its utilization in the
		treatment of various types of hepatic, renal, pulmonary, pancreatic
JOURNAL		and cardiac fibrosis and hypertrophic scars
FEATURES		Patent: WO 0244393-A 3 06-JUN-2002;
source		TGT LAB S A DE C V (MX)
		Location/Qualifiers
		1. .1359
		/organism="Homo sapiens"
		/db_xref="taxon:9606"
BASE COUNT	358 a	374 c 355 g 272 t
ORIGIN		
Query Match	99.4%;	Score 256.4; DB 6; Length 1359;
Best Local Similarity	99.6%;	Pred. No. 1.6e-66;
Matches	257; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
QY	1	AAACCTGCTATGAGGGGAATGCTCACTTTTACCGAGGAAGGCCACACTGACACCATG 60
Db	228	AAACCTGCTATGAGGGGAATGCTCACTTTTACCGAGGAAGGCCACACTGACACCATG 287
QY	61	GGCCGGCCCTGCTGCCCCCTGGAACCTCTGCCACTGTCTTCAGCAACGTAACATGCCAC 120
Db	288	GGCCGGCCCTGCTGCCCCCTGGAACCTCTGCCACTGTCTTCAGCAACGTAACATGCCAC 347
QY	121	AGATCTAATGCTCTTACGCTGGCCCTGGGGAACATTAATTAAGTCAAGAACCCAGACAAC 180
Db	348	AGATCTGATGCTCTTACGCTGGCCCTGGGGAACATTAATTAAGTCAAGAACCCAGACAAC 407
QY	181	CGGAGGCGACCCCTGCTGCTATGTGACAGGTGGGCCCTAAAGCCGCTTGTCCAAGAGTGATG 240
Db	408	CGGAGGCGACCCCTGCTGCTATGTGACAGGTGGGCCCTAAAGCCGCTTGTCCAAGAGTGATG 467
QY	241	GTGCATGACTGCGCAGAT 258
Db	468	GTGCATGACTGCGCAGAT 485
RESULT 15		
E00421		
LOCUS	E00421	1394 bp RNA linear PAT 29-SEP-1997
DEFINITION	CDNA coding urokinase and its flanking region.	
ACCESSION	E00421	
VERSION	E00421.1	GI:2168704
KEYWORDS	JP 1985180591-A/1.	
SOURCE	Homo sapiens.	
ORGANISM	Homo sapiens	
	Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
	1 (bases 1 to 1394)	
	Hiramatsu,T., Kaneda,T., Nagai,M., Arimura,H., Nishida,M. and	
	Suyama,T.	
REFERENCE	DNA SEQUENCE, PLASMID AND HOST OF HUMAN UROKINASE	
AUTHORS	Patent: JP 1985180591-A 1 14-SEP-1985;	
	GREEN CROSS CORP:THE	
	OS Homo sapiens (Human)	
	PN JP 1985180591-A/1	
	PD 14-SEP-1985	
	PF 27-FEB-1984 JP 1984037119	
	PI HIRAMATSU TAKASHI, KANEDA TERUO, NAGAI MASANORI, PI ARIMURA	
	HIROBUMI,	
	PI NISHIDA MASAYUKI, SUYAMA TADAKAZU	
	PC C12N15/00,C07H21/04,C12N1/00//C12N9/72,(C12N1/00,C12N1:19); CC	
	strandedness: Double;	
	CC topology: Linear;	
	CC hypothetical: No;	
	CC anti-sense: No;	
	CC *source: tissue_type=kidney;	
	CC Feature is identified by similarity;	
	FH Key Location/Qualifiers	
	FH	

	FT	sig-peptide	1.	.60	
	FT	mat-peptide	61.	.1296	
	FT		/Product='urokinase'		
	FT	CDS	1.	.1296.	
FEATURES		Location/Qualifiers			
source		1..1394			
		/organism="Homo sapiens"			
		/db_xref="taxon:9606"			
BASE COUNT		349 a	385 c	367 g	293 t
ORIGIN					
Query Match		99.4%;	Score 256.4;	DB 6;	Length 1394;
Best Local Similarity		99.6%;	Pred. No. 1.6e-66;		
Matches 257;	Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	1	AAACCTGCTATGAGGGGAATGTTACTTTTACCGAGGAAGGCCAGCACTGACACCATG	60		
Db	202	AAACCTGCTATGAGGGGAATGTTACTTTTACCGAGGAAGGCCAGCACTGACACCATG	261		
QY	61	GGCCGGCCCTGCTGCCCTGGAACCTCTGCCACTGTCTTCAGCAAAAGTACCATGCCAC	120		
Db	262	GGCCGGCCCTGCTGCCCTGGAACCTCTGCCACTGTCTTCAGCAAAAGTACCATGCCAC	321		
QY	121	AGATCTAATGCTCTTTCAGCTGGGCGCTGGGGAAACATAATTAATCTGCAGGAACCCAGACAAC	180		
Db	322	AGATCTGATGCTCTTTCAGCTGGGCGCTGGGGAAACATAATTAATCTGCAGGAACCCAGACAAC	381		
QY	181	CGAGGCGACCCCTGCTGTGCTATGTCAGGTGGGCCCTAAAGCCGCTGTCCAAAGAGTGCATG	240		
Db	382	CGAGGCGACCCCTGCTGTGCTATGTCAGGTGGGCCCTAAAGCCGCTGTCCAAAGAGTGCATG	441		
QY	241	GTGCATGACTGCGCAGAT	258		
Db	442	GTGCATGACTGCGCAGAT	459		

Search completed: November 11, 2002, 12:24:38
Job time : 2627 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 11, 2002, 11:34:22 ; Search time 231 Seconds
(without alignments)
2515.219 Million cell updates/sec

Title: PCT-US02-27855-2
Perfect score: 258
Sequence: 1 aaacctgtatgaggga.....tgytcatgactgcgcagat 258

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_101002:*

1:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	256.4	99.4	264	24	AAD27075 Human urokinase-ty
2	256.4	99.4	288	24	AAD27083 Human uPA kringlie
3	256.4	99.4	405	24	AAD27078 Human urokinase-ty
4	256.4	99.4	429	24	AAD27082 Human uPA amino te
5	256.4	99.4	1137	11	AAT61672 cDNA encoding huma
6	256.4	99.4	1170	11	AAT61673 Human prourokinase
7	256.4	99.4	1206	11	AAT61674 Human prourokinase
8	256.4	99.4	1212	24	AAD27080 Human urokinase-ty
9	256.4	99.4	1236	11	AAT61671 Human native prour

10	256.4	99.4	1236	11	AAQ04486 Plasmid pUKS3 enco
11	256.4	99.4	1236	11	AAQ06133 Sequence encoding
12	256.4	99.4	1236	11	AAQ06134 Sequence encoding
13	256.4	99.4	1236	11	AAQ06135 Sequence encoding
14	256.4	99.4	1236	12	AAQ10169 Encodes Pro-urokin
15	256.4	99.4	1236	12	AAQ10170 Encodes Pro-urokin
16	256.4	99.4	1236	24	AAD27077 Human urokinase-ty
17	256.4	99.4	1296	11	AAQ06049 plasmid pUK1 pro-u
18	256.4	99.4	1296	15	AAQ55772 Human plasminogen
19	256.4	99.4	1296	24	ABK86598 Pro-urokinase deri
20	256.4	99.4	1473	10	AAQ92037 Sequence of varian
21	256.4	99.4	1474	6	AAQ50138 Sequence encoding
22	256.4	99.4	1475	9	AAQ50141 Sequence encoding
23	256.4	99.4	1475	6	AAQ50158 Pro-UK structural
24	256.4	99.4	1475	20	AAZ24619 Human lung tumor a
25	256.4	99.4	1475	21	AAQ65858 Human lung cancer-
26	256.4	99.4	1475	24	ABL49077 Human lung tumour
27	256.4	99.4	1964	22	AAH28220 Nucleotide sequenc
28	256.4	99.4	1964	24	AAD27855 Human uPA cDNA. H
29	256.4	99.4	2266	10	AAQ93079 Sequence encoding
30	256.4	99.4	2281	20	AAZ24620 Human lung tumor a
31	256.4	99.4	2294	21	AAQ65859 Human lung cancer-
32	256.4	99.4	2294	24	ABL49078 Human lung tumour
33	256.4	99.4	2298	9	AAQ80981 Sequence encoding
34	256.4	99.4	2299	7	AAQ60703 Sequence encoding
35	256.4	99.4	2301	8	AAQ70390 cDNA encoding natu
36	256.4	99.4	2301	10	AAQ91075 DNA encoding natu
37	256.4	99.4	2303	10	AAQ91740 Sequence of prouro
38	256.4	99.4	2303	15	AAQ73483 Full length human
39	256.4	99.4	2304	4	AAQ30030 Sequence encoding
40	256.4	99.4	2304	8	AAQ71338 Modified prourokin
41	256.4	99.4	2304	8	AAQ71368 Modified prourokin
42	256.4	99.4	2304	8	AAQ71369 Modified prourokin
43	256.4	99.4	2304	24	ABL59543 Human pro-urokinas
44	256.4	99.4	2377	13	AAQ20360 Human pro-urokinas
45	256.4	99.4	2427	11	AAQ04107 Human pro-urokinas

ALIGNMENTS

RESULT 1	
AAD27075	
ID	AAD27075 standard; DNA; 264 BP.
AC	AAD27075;
XX	
DT	09-APR-2002 (first entry)
XX	
DE	Human urokinase-type plasminogen activator (uPA) kringlie DNA.
XX	
KW	Human; urokinase-type plasminogen activator; uPA; therapy; hypertension; stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder; microvascular occlusion; angioemic disorder; pulmonary fibrosis; asthma; tumour cell metastasis; glioma; diabetic retinopathy; wound healing; clotting disorder; uterine contraction disorder; respiratory disease; male impotence; adult respiratory distress syndrome; ds.
KW	
XX	
OS	Homo sapiens.
XX	
FH	
FT	Key
FT	CDS
FT	Location/Qualifiers
FT	1..264
FT	/*tag= a
FT	/product= "Human uPA kringlie"
FT	/note= "CDS does not include start and stop codon"
FT	/partial
PN	WO200197752-A2.
XX	
PD	27-DEC-2001.
XX	
PF	13-JUN-2001; 2001WO-US18976.
XX	

occlusions, thrombotic microangiopathies, surgically induced thrombotic disorders, angiotonic disorders, pulmonary fibrosis, asthma, tumour cell invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma, diabetic retinopathy, wound healing, clotting disorder, uterine contraction disorder, male impotence, respiratory disease or condition such as asthma, adult respiratory distress syndrome, primary pulmonary hypertension, microvascular thrombotic occlusion, and a disorder associated with chronic intrapulmonary fibrin formation. The present sequence is a DNA encoding human urokinase-type plasminogen activator (uPA) amino terminal fragment (ATF) and connecting peptide.

Sequence 429 BP; 120 A; 115 C; 107 G; 87 T; 0 other;

Query Match 99.4%; Score 256.4; DB 24; Length 429;
Best Local Similarity 99.6%; Pred. No. 1.2e-34;
Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAACCTGCTATGAGGGGATGCTCACTTTTACCAGGAAAGCCAGCACTGACACCATG 60
DB 142 AAAACCTGCTATGAGGGGATGCTCACTTTTACCAGGAAAGCCAGCACTGACACCATG 201
QY 61 GGCCGGCCCTGCTGCTGCTGCACTGCTGCTTTCAGCAACGATGACATGCCCAC 120
DB 202 GGCCGGCCCTGCTGCTGCTGCACTGCTGCTTTCAGCAACGATGACATGCCCAC 261
QY 121 AGATCTAATGCTCTTCACTGCGCTGGGAAACATTAATTTACTGAGGAACCCAGACAAC 180
DB 262 AGATCTGATGCTCTTCACTGCGCTGGGAAACATTAATTTACTGAGGAACCCAGACAAC 321
QY 181 CGGAGCGACCCCTGCTGCTATGTCAGAGTGGGCTTAAGCCGCTTGTCCAAGAGTGCATG 240
DB 322 CGGAGCGACCCCTGCTGCTATGTCAGAGTGGGCTTAAGCCGCTTGTCCAAGAGTGCATG 381
QY 241 GTGCATGACTGCGCAGAT 258
DB 382 GTGCATGACTGCGCAGAT 399

RESULT 5
AAT61672
ID AAT61672 standard; cDNA; 1137 BP.

AC AAT61672;

DT 04-JUN-1997 (first entry)

DE cDNA encoding human prourokinase variant lacking entire EGF domain.

XX Human; prourokinase; hPUK; variant; half-life; increase; EGF;

KW epidermal growth factor domain; deletion; thrombolysis;

KW fibrinolysis; ds.

OS Homo sapiens.

OS Synthetic.

XX Key CDS Location/Qualifiers

FT 1..1137

FT /*tag= a

FT /product= deltaE1E2E3-PUK

FT /transl_except= pos:430..432, aa:Thr

FT /note= "encodes human PUK lacking the entire EGF domain"

FT mutation 27..28

FT /*tag= b

FT /note= "region encoding entire EGF domain has been deleted from between these positions"

XX EP398361-A.

XX 22-NOV-1990.

XX 18-MAY-1990; 90EP-0109472.

PR 22-FEB-1990; 90JP-0042020.
PR 18-MAY-1989; 89JP-0126433.
PR 03-JUL-1986; 86JP-0156936.
PR 18-FEB-1987; 87JP-0036495.
PR 18-MAY-1989; 89JP-0126434.

XX (GREG) GREEN CROSS CORP.

XX Airmura H, Amatsuji Y, Hirose M, Kasai S, Kawabe H;

PI Morita M, Tanabe T;

XX WPI; 1990-350146/47.

DR P-PSDB; AAW13635.

PT Human pro-urokinase variants - deficient in loop regions of

PT epidermal growth factor, showing long blood half-life, as

PT fibrinolytic agent

PS Claim 2; Page -; 22pp; English.

CC New variants of human prourokinase (hPUK) comprise a hPUK deficient in (i) at least part of the first loop region of the epidermal growth factor (EGF) domain; (ii) at least part of the first loop and at least part of the second loop; or (iii) at least part of the third loop. The hPUK variants show an increased blood half-life comparable to that of the whole EGF domain-deficient hPUK variant and urokinase while retaining the same properties as those of hPUK. They have potent thrombolytic activity and very little tendency to cause spontaneous bleeding. The present sequence encodes a specific variant of hPUK which lacks the entire EGF domain; the sequence does not appear in the specification and has been created using the wild-type coding sequence and the junction sequence after deletion, both of which are given (in Fig 1 and in Fig 2(3), respectively).

CC Sequence 1137 BP; 287 A; 319 C; 297 G; 234 T; 0 other;

Query Match 99.4%; Score 256.4; DB 11; Length 1137;
Best Local Similarity 99.6%; Pred. No. 9.4e-35;
Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAACCTGCTATGAGGGGATGCTCACTTTTACCAGGAAAGCCAGCACTGACACCATG 60
DB 43 AAAACCTGCTATGAGGGGATGCTCACTTTTACCAGGAAAGCCAGCACTGACACCATG 102

QY 61 GGCCGGCCCTGCTGCTGCTGCACTGCTGCTTTCAGCAACGATGACATGCCCAC 120
DB 103 GGCCGGCCCTGCTGCTGCTGCACTGCTGCTTTCAGCAACGATGACATGCCCAC 162

QY 121 AGATCTAATGCTCTTCACTGCGCTGGGAAACATTAATTTACTGAGGAACCCAGACAAC 180
DB 163 AGATCTGATGCTCTTCACTGCGCTGGGAAACATTAATTTACTGAGGAACCCAGACAAC 222

QY 181 CGGAGCGACCCCTGCTGCTATGTCAGAGTGGGCTTAAGCCGCTTGTCCAAGAGTGCATG 240
DB 223 CGGAGCGACCCCTGCTGCTATGTCAGAGTGGGCTTAAGCCGCTTGTCCAAGAGTGCATG 282

QY 241 GTGCATGACTGCGCAGAT 258
DB 283 GTGCATGACTGCGCAGAT 300

RESULT 6

AAT61673
ID AAT61673 standard; cDNA; 1170 BP.

AC AAT61673;

DT 04-JUN-1997 (first entry)

DE Human prourokinase variant lacking EGF domain loops 1 and 2, cDNA.

XX Human; prourokinase; hPUK; variant; half-life; increase; EGF;

XX Human; prourokinase; hPUK; variant; half-life; increase; EGF;

XX Human; prourokinase; hPUK; variant; half-life; increase; EGF;

XX Human; prourokinase; hPUK; variant; half-life; increase; EGF;

XX Human; prourokinase; hPUK; variant; half-life; increase; EGF;

XX Human; prourokinase; hPUK; variant; half-life; increase; EGF;

XX Human; prourokinase; hPUK; variant; half-life; increase; EGF;

KM		epidermal growth factor domain; deletion; thrombolysis;
KW		fibrinolysis; ds.
XX		
OS	Homo sapiens.	
OS	Synthetic.	
XX		
FH		
FT	Key	Location/Qualifiers
FT	CDS	1..1170
FT		/tag- a
FT		/product= deltaE1E2-PUK
FT		/transl_except= pos:463..465, aa:Thr
FT		/note= "encodes human PUK lacking loops 1 and 2 of the EGF domain"
FT	mutation	30..31
FT		/tag- b
FT		/note= "codons 11-32 have been deleted from between these positions"
PN		
XX	EP398361-A.	
XX		
PD	22-NOV-1990.	
XX		
PF	18-MAY-1990;	90EP-0109472.
XX		
PR	22-FEB-1990;	90JP-0042020.
PR	18-MAY-1989;	89JP-0126433.
PR	03-JUL-1986;	86JP-0156936.
PR	18-FEB-1987;	87JP-0036495.
PR	18-MAY-1989;	89JP-0126434.
XX		
PA	(GREC) GREEN CROSS CORP.	
PI	Airmura H, Amatsuji Y, Hirose M, Kasai S, Kawabe H;	
PI	Morita M, Tanabe T;	
XX		
DR	WPI; 1990-350146/47.	
P-PSDB;	AAW13636.	
.		
PT	Human pro-urolkinase variants - deficient in loop regions of	
PT	epidermal growth factor, showing long blood half-life, as	
PT	fibrinolytic agent	
XX		
PS	Claim 7; Page -: 22pp; English.	
XX		
CC	New variants of human prourokinase (hPUK) comprise a hPUK deficient	
CC	ln (I) at least part of the first loop region of the epidermal growth	
CC	factor (EGF) domain; (II) at least part of the first loop and at	
CC	least part of the second loop; or (III) at least part of the third	
CC	loop. The hPUK variants show an increased blood half-life comparable	
CC	to that of the whole EGF domain-deficient hPUK variant and urokinase	
CC	while retaining the same properties as those of hPUK. They have	
CC	potent thrombolytic activity and very little tendency to cause	
CC	spontaneous bleeding. The present sequence encodes a specific variant	
CC	of hPUK which lacks loops 1 and 2 of the EGF domain; the sequence does	
CC	not appear in the specification and has been created using the	
CC	wild-type coding sequence and the junction sequence after	
CC	deletion, both of which are given (in Fig 1 and on page 8,	
CC	respectively).	
XX		
SQ	Sequence 1170 BP; 299 A; 326 C; 306 G; 239 T; 0 other;	
	Query Match	99.4%; Score 256.4; DB 11; Length 1170;
	Best Local Similarity	99.6%; Pred. No. 9.4e-35;
	Matches 257; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
OY	1 AAACCTGCTATGAGGGGAATGTCACTTTTACCAGGAAGGCCAGCACTGCACCATG 60	
Db	76 AAAACCTGCTATGAGGGGAATGTCACTTTTACCAGGAAGGCCAGCACTGCACCATG 135	
OY	61 GGCGGGCCCTGCTGCCCCCTGGAACTCTGCCACTGTCTTCAGCAAACGTACCATGCCCAC 120	
Db	136 GGCGGGCCCTGCTGCCCCCTGGAACTCTGCCACTGTCTTCAGCAAACGTACCATGCCCAC 195	

OY		121	AGATCTAATGCTCTTCACGTGGCGCCCTGGGGAAACATAATTACTGCAGGAACCCAGACAAC	180
Dd		196	AGATCTGATGCTCTTCACGTGGCGCCCTGGGGAAACATAATTACTGCAGGAACCCAGACAAC	255
OY		181	CGGAGCGCACCCCTGCTGCTATGTGCAGGTGGGCCCTAAAGCCGCTTGTCCAAGAGTGCATG	240
Dd		256	CGGAGCGCACCCCTGCTGCTATGTGCAGGTGGGCCCTAAAGCCGCTTGTCCAAGAGTGCATG	315
OY		241	GTGCATGACTGCCGAGAT	258
Dd		316	GTGCATGACTGCCGAGAT	333
RESULT 7				
AAT61674				
ID	AAT61674	standard;	CDNA; 1206 BP.	
XX				
AC	AAT61674;			
XX				
DT	04-JUN-1997	(first entry)		
XX				
DE	Human prourokinase variant lacking EGF domain loop 3, cDNA.			
XX				
KW	Human; prourokinase; hPUK; variant; half-life; increase; EGF;			
KM	epidermal growth factor domain; deletion; thrombolysis;			
KW	fibrinolysis; ds.			
XX				
OS	Homo sapiens.			
OS	Synthetic.			
XX				
FH	Key	Location/Qualifiers		
FT	CDS	1..1206		
FT		/tag= a		
FT		/product= deltaE3-PUK		
FT		/transl_except= pos:499..501, aa:Thr		
FT		/note= "encodes human PUK lacking loop 3		
FT		of the EGF domain"		
mutation		96..97		
FT		/tag= b		
FT		/note= "codons 33-42 have been deleted from between		
FT		these positions"		
XX				
PX	EP398361-A.			
XX				
PD	22-NOV-1990.			
XX				
PF	18-MAY-1990;	90EP-0109472.		
XX				
PR	22-FEB-1990;	90JP-0042020.		
PR	18-MAY-1989;	89JP-0126433.		
PR	03-JUL-1986;	86JP-0156936.		
PR	18-FEB-1987;	87JP-0036495.		
PR	18-MAY-1989;	89JP-0126434.		
XX				
PA	(GREC) GREEN CROSS CORP.			
PI	Airmura H, Amatsuji Y, Hirose M, Kasai S, Kawabe H;			
PI	Morita M, Tanabe T;			
XX				
DR	WPI; 1990-350146/47.			
DR	P-PSDB; AAWI3637.			
XX				
PT	Human pro-uroporphyrinogen decarboxylase - deficient in liver regions of			
PT	epidermal growth factor, showing long blood half-life, as			
PT	fibrinolytic agent			
XX				
PX	Claim 12; Page -; 22pp; English.			
XX				
CC	New variants of human prourokinase (hPUK) comprise a hPUK deficient			
CC	in (i) at least part of the first loop region of the epidermal growth			
CC	factor (EGF) domain; (ii) at least part of the first loop and at			
CC	least part of the second loop; or (iii) at least part of the third			
CC	loop. The hPUK variants show an increased blood half-life comparable			


```
FH Key Location/Qualifiers
FT CDS 1..1236
FT /*tag= a
FT /product= prourokinase
FT /transl_except= pos:528..530, aa:Thr
FT /note= "the EGF domain is encoded by nucleotides
FT 28-126; novel variants of PUK which lack
FT at least part of the EGF domain are claimed"
FT Misc-difference 177
FT /note= "corresponds to TAC codon"
XX
XX PN EP398361-A.
XX PD 22-NOV-1990.
XX PF 18-MAY-1990; 90EP-0109472.
XX PR 22-FEB-1990; 90JP-0042020.
XX PR 18-MAY-1989; 89JP-0126433.
XX PR 03-JUL-1986; 86JP-0156936.
XX PR 18-FEB-1987; 87JP-0036495.
XX PR 18-MAY-1989; 89JP-0126434.
XX PA (GREC ) GREEN CROSS CORP.
XX PI Alimura H, Amatsuji Y, Hirose M, Kasai S, Kawabe H;
XX PI Morita M, Tanabe T;
XX DR WPI; 1990-350146/47.
XX DR P-PSDB; AAW13634.
XX PT Human pro-urokinase variants - deficient in loop regions of
XX PT epidermal growth factor, showing long blood half-life, as
XX PT fibrinolytic agent
XX PS Disclosure; Fig 1; 22pp; English.
XX CC New variants of human prourokinase (hPUK) comprise a hPUK deficient
XX CC in (1) at least part of the first loop region of the epidermal growth
XX CC factor (EGF) domain; (11) at least part of the first loop and at
XX CC least part of the second loop; or (111) at least part of the third
XX CC loop. The hPUK variants show an increased blood half-life comparable
XX CC to that of the whole EGF domain-deficient hPUK variant and urokinase
XX CC while retaining the same properties as those of hPUK. They have
XX CC potent thrombolytic activity and very little tendency to cause
XX CC spontaneous bleeding. The present sequence encodes the wild-type
XX CC hPUK protein, including the EGF domain.
XX SQ Sequence 1236 BP; 318 A; 341 C; 320 G; 257 T; 0 other;
SQ
Query Match 99.4%; Score 256.4; DB 11; Length 1236;
Best Local Similarity 99.6%; Pred. No. 9.2e-35;
Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AAAACCTGCTATGAGGGGAATGTCACCTTTTACCGAGGAAGCCAGCACTGACACCATG 60
DB 142 AAAACCTGCTATGAGGGGAATGTCACCTTTTACCGAGGAAGCCAGCACTGACACCATG 201
QY 61 GGCCGGCCCTGCTGCTGCACTCTGCCACTGTCTTCAGCAAAAGCTACCATGCCAC 120
DB 202 GGCCGGCCCTGCTGCTGCACTCTGCCACTGTCTTCAGCAAAAGCTACCATGCCAC 261
QY 121 AGATCTAATGCTCTTCAGCTGGGCTGGGGAACATATTAATCTGCAAGAACCCAGACAAC 180
DB 262 AGATCTAATGCTCTTCAGCTGGGCTGGGGAACATATTAATCTGCAAGAACCCAGACAAC 321
QY 181 CGGAGGCAACCTGCTGCTATGTCAGAGGTGGGCTTAAGCCGCTTGTCCAAGAGTGCATG 240
DB 322 CGGAGGCAACCTGCTGCTATGTCAGAGGTGGGCTTAAGCCGCTTGTCCAAGAGTGCATG 381
QY 241 GTGCATGACTGCGCAGAT 258
DB 382 GTGCATGACTGCGCAGAT 399
```

```
RESULT 10
AAQ04486
ID AAQ04486 standard; DNA; 1236 BP.
XX AC AAQ04486;
XX DT 04-OCT-1990 (first entry)
XX DE Plasmid pUKS3 encoding UK-S3.
XX KW Urokinase; glycosylation; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT misc_difference 457.459
FT /*tag= a
FT /label=synthetic_mutation
FT /note="old seq (ctg)"
FT misc_difference 463.465
FT /*tag= b
FT /label=synthetic_mutation
FT /note="old seq (ccc)"
XX PN EP370205-A.
XX PD 30-MAY-1990.
XX PF 28-SEP-1989; 89EP-0117981.
XX PR 29-SEP-1989; 89JP-0245705.
XX PA (KYOW ) KYOWA HAKKA KK.
XX PI Sasaki K, Nishi T, Yasumuru S, Sato M, Itoh S;
XX DR WPI; 1990-165029/22.
XX DR P-PSDB; AAR05117.
XX PT Polypeptide(s) with added carbohydrate chains - formed by
XX PT modification of amino acid sequence, used to improve
XX PS physicochemical properties and/or activities.
XX PS Disclosure; ; 30pp; English.
XX CC The sequence encodes a deriv. of mature urokinase, designated
XX CC UK-S3 which has 2 amino acid substns. which result in an N-linked
XX CC glycosylation site giving the new protein improved stability and
XX CC activity.
XX CC See also AAQ04481-85.
XX SQ Sequence 1236 BP; 325 A; 336 C; 316 G; 259 T; 0 other;
SQ
Query Match 99.4%; Score 256.4; DB 11; Length 1236;
Best Local Similarity 99.6%; Pred. No. 9.2e-35;
Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AAAACCTGCTATGAGGGGAATGTCACCTTTTACCGAGGAAGCCAGCACTGACACCATG 60
DB 142 AAAACCTGCTATGAGGGGAATGTCACCTTTTACCGAGGAAGCCAGCACTGACACCATG 201
QY 61 GGCCGGCCCTGCTGCTGCACTCTGCCACTGTCTTCAGCAAAAGCTACCATGCCAC 120
DB 202 GGCCGGCCCTGCTGCTGCACTCTGCCACTGTCTTCAGCAAAAGCTACCATGCCAC 261
QY 121 AGATCTAATGCTCTTCAGCTGGGCTGGGGAACATATTAATCTGCAAGAACCCAGACAAC 180
DB 262 AGATCTAATGCTCTTCAGCTGGGCTGGGGAACATATTAATCTGCAAGAACCCAGACAAC 321
QY 181 CGGAGGCAACCTGCTGCTATGTCAGAGGTGGGCTTAAGCCGCTTGTCCAAGAGTGCATG 240
```


Db 322 CGGAGCGGACCCCTGGTCTATGTGCAGGTGGGCGCTAAAGCCGCTGTGCCAAGAGTGCATG 381
QY 241 GTGCATGACTGGCGAGAT 258
|||||
Db 382 GTGCATGACTGGCGAGAT 399

RESULT 13

AAQ06135
ID AAQ06135 standard; DNA; 1236 BP.

XX
AC AAQ06135;

DT 21-FEB-1991 (first entry)

DE Sequence encoding human pro-urokinase variant.

KW Thrombin; fibrin; bleeding; pHR27; ds.

OS Homo sapiens.

Key Location/Qualifiers

FT CDS 1..1233

FT /tag= a

PN EP398362-A.

PD 22-NOV-1990.

PF 18-MAY-1990; 90EP-0109473.

PR 18-MAY-1989; 89JP-0126434.

PA (GREG) GREEN CROSS CORP.

PI Tanabe T, Amatsuji Y, Kasai S, Hirose M, Morita M;

PI Kawabe H, Arimura H;

DR WPI; 1990-350147/47.

DR P-PSDB; AAR07904.

PT Human pro-urokinase variant - produced by recombinant methods,

PS showing increased half life in blood and high affinity for fibrin.

CC Disclosure; Fig 1; 27pp; English.

CC Modified pro-urokinase has a longer half-life in blood, and dissolves

CC thrombin without causing the spontaneous bleeding associated with

CC urokinase. The modification puts an epidermal growth factor domain

CC into the protein, containing the sequence Asn-X-Ser or Asn-X-Thr

CC where X is any residue.

CC Plasmid pHR27 is disclosed as containing the modified sequence.

XX Sequence 1236 BP; 317 A; 342 C; 319 G; 258 T; 0 other;

QY 1 AAAACCTGCTATGAGGGGAATGTCACCTTTTACCGAGGAAGGCCAGACTGACACCATG 60

Db 142 AAAACCTGCTATGAGGGGAATGTCACCTTTTACCGAGGAAGGCCAGACTGACACCATG 201

QY 61 GGCCGGCCCTGCGCTGCGCTGGAAGTCTGCCACTGTCTTCAGCAAAAGCTACATGCCAC 120

Db 202 GGCCGGCCCTGCGCTGCGCTGGAAGTCTGCCACTGTCTTCAGCAAAAGCTACATGCCAC 261

QY 121 AGATCTAATGCTCTTACAGCTGGGCTGGGGAACAATAATTACTGCAGGAAGCCAGACAAC 180

Db 262 AGATCTAATGCTCTTACAGCTGGGCTGGGGAACAATAATTACTGCAGGAAGCCAGACAAC 321

QY 181 CGGAGCGGACCCCTGGTCTATGTGCAGGTGGGCGCTAAAGCCGCTGTGCCAAGAGTGCATG 240

|||||

Db 322 CGGAGCGGACCCCTGGTCTATGTGCAGGTGGGCGCTAAAGCCGCTGTGCCAAGAGTGCATG 381
QY 241 GTGCATGACTGGCGAGAT 258
|||||
Db 382 GTGCATGACTGGCGAGAT 399

RESULT 14

AAQ10169
ID AAQ10169 standard; DNA; 1236 BP.

XX
AC AAQ10169;

DT 18-MAR-1991 (first entry)

DE Encodes Pro-urokinase derivative UK-T4 with Ser(153) and Thr(155).

KW pro-urokinase; UK-T4; plasminogen activator; myocardial infarction;

OS cerebral thrombosis; ss.

OS Homo sapiens.

Key Location/Qualifiers

FT CDS 1..1236

FT /tag= a

FT /product= UK-T4

PN EP405285-A.

PD 02-JAN-1991.

PF 18-JUN-1990; 90EP-0111471.

PR 19-JUN-1989; 89JP-0156302.

PA (KYOW) KYOWA HAKKO KOGYO KK.

PI Yasamura S, Nishi T, Ito S;

DR WPI; 1991-008678/02.

DR P-PSDB; AAR10057.

PT New plasminogen activator almost identical to natural

PT pro-urokinase - is thrombin resistant and used for

PT prophylaxis, treatment of cerebral thrombosis or myocardial

PT infarction

PS Disclosure; Page 8; 84pp; English.

CC UK-T4 is one example of a plasminogen activator which differs from

CC natural human pro-urokinase at positions 153 and 155. There are

CC corresponding codon changes at position 457-459 (Leu substituted by

CC Ser) and at position 463 to 465 (Pro substituted by Thr) of this

CC DNA sequence, relative to the wild-type coding region. pHPA2,

CC constructed from pUK1, is cleaved with EcoRI and HindIII and a

CC 3.4kb DNA fragment is obtained. pUKS1 is cleaved with HindIII and

CC CfrI to give a 0.75kb fragment. The two fragments, together with

CC two phosphorylated synthetic DNAs for site-specific mutation, were

CC ligated together in the presence of T4 ligase to give pUK-T4 which

CC has the sequence given here. See also AAQ10168 and AAQ10170.

XX Sequence 1236 BP; 321 A; 338 C; 318 G; 259 T; 0 other;

QY 61 GGCCGGCCCTGCGCTGCGCTGGAAGTCTGCCACTGTCTTCAGCAAAAGCTACATGCCAC 120

|||||

Db 202 GGCCGCCCCCTGCTCCCTGGAACCTCTGCCACTGTCCCTTCAGCAAAACGTACCATGCCAC 261
QY 121 AGATCTAATGCTCTTTCAGCTGGGCTGGGAAACATAATTAATCTGCAGAACCCAGACAAC 180
Db 262 AGATCTGATGCTCTTTCAGCTGGGCTGGGAAACATAATTAATCTGCAGAACCCAGACAAC 321
QY 181 CGGAGCGGACCCCTGCTGCTATGTGAGAGTGGGCTTAAAGCCGCTTGTCCAGAGAGTGATG 240
Db 322 CGGAGCGGACCCCTGCTGCTATGTGAGAGTGGGCTTAAAGCCGCTTGTCCAGAGAGTGATG 381
QY 241 GTGCATGACTGCGCAGAT 258
Db 382 GTGCATGACTGCGCAGAT 399

RESULT 15
AAQ10170
ID AAQ10170 standard; DNA; 1236 BP.
XX
AC AAQ10170;
XX
DT 18-MAR-1991 (first entry)
XX
DE Encodes Pro-urokinase derivative UK-S3 with Asn(153) and Thr(155).
XX
KW pro-urokinase; UK-S3; plasminogen activator; myocardial infarction;
KW cerebral thrombosis; ss.
XX
OS Homo sapiens.

XX
FH Key Location/Qualifiers
FT CDS 1..1236
FT /*tag= a
FT /product= UK-S3

PN EP405285-A.
XX
PD 02-JAN-1991.
XX
PF 18-JUN-1990; 90EP-0111471.
XX
PR 19-JUN-1989; 89JP-0156302.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI Yasamura S, Nishi T, Ito S;
XX
DR WPI; 1991-008678/02.
DR P-PSDB; AAR10058.

XX
PT New plasminogen activator almost identical to natural
PT pro:urokinase - is thrombin resistant and used for
PT prophylaxis, treatment of cerebral thrombosis or myocardial
PT infarction

XX
PS Disclosure; Page 9; 84pp; English.

XX
CC UK-S3 is one example of a plasminogen activator which differs from
CC natural human pro-urokinase at positions 153 and 155. There are
CC corresponding codon changes at position 457-459 (Leu substituted by
CC Asn) and at position 463 to 465 (Pro substituted by Thr) of this
CC DNA sequence, relative to the wild-type coding region. phpA2,
CC constructed from pUK1, is cleaved with EcoRI and HindIII and a
CC 3.4kb DNA fragment is obtained. pUKS1 is cleaved with HindIII and
CC CfrI to give a 0.75kb fragment. The two fragments, together with
CC two phosphorylated synthetic DNAs for site-specific mutation, were
CC ligated together in the presence of T4 ligase to give pUK-S3 which
CC has the sequence given here. See also AAQ10168 and AAQ10169.
XX

SQ Sequence 1236 BP; 322 A; 338 C; 317 G; 259 T; 0 other;

Query Match 99.4%; Score 256.4; DB 12; Length 1236;
Best Local Similarity 99.6%; Pred. No. 9.2e-35;

Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AAACCTGCTATGAGGGGAATGGTCACTTTTACCAGAGAAAGGCCACCTGACACCATG 60
Db 142 AAACCTGCTATGAGGGGAATGGTCACTTTTACCAGAGAAAGGCCACCTGACACCATG 201
QY 61 GGCCGCCCCCTGCTGCCCCCTGGAACCTGTGCCACTGTCTTCCAGCAAAACGTACCATGCCAC 120
Db 202 GGCCGCCCCCTGCTGCCCCCTGGAACCTGTGCCACTGTCTTCCAGCAAAACGTACCATGCCAC 261
QY 121 AGATCTAATGCTCTTTCAGCTGGGCTGGGAAACATAATTAATCTGCAGAGAACCCAGACAAC 180
Db 262 AGATCTGATGCTCTTTCAGCTGGGCTGGGAAACATAATTAATCTGCAGAGAACCCAGACAAC 321
QY 181 CGGAGCGGACCCCTGCTGCTATGTGAGAGTGGGCTTAAAGCCGCTTGTCCAGAGAGTGATG 240
Db 322 CGGAGCGGACCCCTGCTGCTATGTGAGAGTGGGCTTAAAGCCGCTTGTCCAGAGAGTGATG 381
QY 241 GTGCATGACTGCGCAGAT 258
Db 382 GTGCATGACTGCGCAGAT 399

Search completed: November 11, 2002, 11:40:39
Job time : 234 secs


```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 423 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..419
; SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-984-186-11

```

Query Match	99.48%	Score 256.4;	DB 10;	Length 423;
Best Local Similarity	99.68%	Pred. No. 4.9e-77;		
Matches 257; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

QY	1	AAACCTGCTATGAGGGGAATGGTCACTTTTACCAGAGAAAGGCCAGCAGTGCACCATG	60
Db	153	AAAACCTGTATGAGGGGAATGGTCACTTTTACCAGAGAAAGGCCAGCAGTGCACCATG	212
QY	61	GGCCGGCCCTGCGCTGCCCTGGAACCTCTGCCACTGTCTTCAGCAAACGTACCATGCCAC	120
Db	213	GGCCGGCCCTGCGCTGCCCTGGAACCTCTGCCACTGTCTTCAGCAAACGTACCATGCCAC	272
QY	121	AGATCTAATGCTCTTTCAGCTGGGCGCTGGGGAACAATATTACTGCAGAACCCAGACAAC	180
Db	273	AGATCTAATGCTCTTTCAGCTGGGCGCTGGGGAACAATATTACTGCAGAACCCAGACAAC	332
QY	181	CGGAGGCGAACCCTGGTGCTATGTGCAGAGTGGGCCCTAAAGCCGCTGTCCAAGAGTGCATG	240
Db	333	CGGAGGCGAACCCTGGTGCTATGTGCAGAGTGGGCCCTAAAGCCGCTGTCCAAGAGTGCATG	392
QY	- 241	GTCGATGACTGCCGAGAT	258
Db	393	GTCGATGACTGCCGAGAT	410

```

RESULT 5
US-09-880-503-17
: Sequence 17, Application US/09880503
: Patent No. US20020131964A1
: GENERAL INFORMATION:
: APPLICANT: CINES, Douglas B
: APPLICANT: HIGAZI, Abd Al-Roof
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
: TITLE OF INVENTION: TISSUE CONTRACTABILITY
: FILE REFERENCE: 9596-331
: CURRENT APPLICATION NUMBER: US/09/880,503
: CURRENT FILING DATE: 2001-06-13
: PRIOR APPLICATION NUMBER: US 60/212,847
: PRIOR FILING DATE: 2000-06-20
: NUMBER OF SEQ ID NOS: 18
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 17
: LENGTH: 429
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-880-503-17

```

Query Match	99.48;	Score 256.4;	DB 10;	Length 429;
Best Local Similarity	99.68;	Pred. No. 5e-77;		
Matches 257; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

[illegible]

Db	262	AGATCTGATGCTCTTCACAGCTGGGGCCCTGGGGGAAACATTAATTACTGCAGGAACCCAGACAAAC	321
QY	181	CGGAGGCGACCCCTGGTGTCTATGTGACAGTGGGGCCTAAAGCCGCTTGTCCAAAGAGTGCATG	240
Db	322	CGGAGGCGACCCCTGGTGTCTATGTGACAGTGGGGCCTAAAGCCGCTTGTCCAAAGAGTGCATG	381
QY	241	GTGCATGACTGCGCAGAT	258
Db	382	GTGCATGACTGCGCAGAT	399

```

RESULT 6
US-09-880-503-15
; Sequence 15, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 1212
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-880-503-15

```

Query Match	99.48;	Score 256.4;	DB 10;	Length 1212;
Best Local Similarity	99.68;	Pred. No. 7.9e-77;		
Matches 257; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

QY	1	AAACCTGCTATGAGGGGAATGGTCACTTTTACCAGGAAGGCCAGCAGCTGCACCATG	60
Db	142	AAAACCTGTATGAGGGGAATGGTCACTTTTACCAGGAAGGCCAGCAGCTGCACCATG	2011
QY	61	GGCCGGCCCTGCGCTGCCCTGGAACCTGCCCAGTCTTCAGCAAACGTACCATGCCAC	120
Db	202	GCGCCGCCCTGCGCTGCCCTGGAACCTGCCCAGTCTTCAGCAAACGTACCATGCCAC	261
QY	121	AGATCTAATGCTCTTCAGCTGGGCGTGGGGAACAATAATTACTGCAGAACCAGACAAC	180
Db	262	AGATCTGATGCTCTTCAGCTGGGCGTGGGGAACAATAATTACTGCAGAACCAGACAAC	321
QY	181	CGGAGGCGAACCCCTGGTGCTATGTGACGAGTGGGCGCTAAAGCCGCTTGTCCAAGAGTGCATG	240
Db	322	CGGAGGCGAACCCCTGGTGCTATGTGACGAGTGGGCGCTAAAGCCGCTTGTCCAAGAGTGCATG	381
QY	241	GTGCATGACTGCCAGAT	258
Db	382	GTGCATGACTGCCAGAT	399

```

RESULT 7
US-09-880-503-12
; Sequence 12, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18

```


SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 1236
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-880-503-12

Query Match 99.4%; Score 256.4; DB 10; Length 1236;
Best Local Similarity 99.6%; Pred. No. 7.9e-77;
Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAACCTGCTATGAGGGGAATGGTCACTTTTACCGAGAAAGGCCAGCACTGACACCATG 60
DB 142 AAAACCTGCTATGAGGGGAATGGTCACTTTTACCGAGAAAGGCCAGCACTGACACCATG 201
QY 61 GGCCGGCCCTGCGCTGCCCTGGAACCTCTGCCACTGTCTTCAGCAAAAGCTACCATGCCAC 120
DB 202 GGCCGGCCCTGCGCTGCCCTGGAACCTCTGCCACTGTCTTCAGCAAAAGCTACCATGCCAC 261
QY 121 AGATCTAATGCTCTTCAAGCTGGGCGGAAACATAATTAATCTGAGGAAGCCAGACAAC 180
DB 262 AGATCTGATGCTCTTCAAGCTGGGCGGAAACATAATTAATCTGAGGAAGCCAGACAAC 321
QY 181 CGGAGCGGACCCCTGGTGTCTATGTGAGAGTGGGCGCTAAAGCCGCTTGTCCAAGAGTGCATG 240
DB 322 CGGAGCGGACCCCTGGTGTCTATGTGAGAGTGGGCGCTAAAGCCGCTTGTCCAAGAGTGCATG 381
QY 241 GTGCATGACTGCGCAGAT 258
DB 382 GTGCATGACTGCGCAGAT 399

RESULT 8

US-09-735-705-122
; Sequence 122, Application US/09735705
; Patent No. US20020052329A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skelky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C14
; CURRENT APPLICATION NUMBER: US/09/735,705
; CURRENT FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 419
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 122
; LENGTH: 1475
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-735-705-122

Query Match 99.4%; Score 256.4; DB 10; Length 1475;
Best Local Similarity 99.6%; Pred. No. 8.6e-77;
Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAACCTGCTATGAGGGGAATGGTCACTTTTACCGAGAAAGGCCAGCACTGACACCATG 60
DB 282 AAAACCTGCTATGAGGGGAATGGTCACTTTTACCGAGAAAGGCCAGCACTGACACCATG 341
QY 61 GGCCGGCCCTGCGCTGCCCTGGAACCTCTGCCACTGTCTTCAGCAAAAGCTACCATGCCAC 120
DB 342 GGCCGGCCCTGCGCTGCCCTGGAACCTCTGCCACTGTCTTCAGCAAAAGCTACCATGCCAC 401

QY 121 AGATCTAATGCTCTTCAAGCTGGGCGGAAACATAATTAATCTGAGGAAGCCAGACAAC 180
DB 402 AGATCTGATGCTCTTCAAGCTGGGCGGAAACATAATTAATCTGAGGAAGCCAGACAAC 461
QY 181 CGGAGCGGACCCCTGGTGTCTATGTGAGAGTGGGCGCTAAAGCCGCTTGTCCAAGAGTGCATG 240
DB 462 CGGAGCGGACCCCTGGTGTCTATGTGAGAGTGGGCGCTAAAGCCGCTTGTCCAAGAGTGCATG 521
QY 241 GTGCATGACTGCGCAGAT 258
DB 522 GTGCATGACTGCGCAGAT 539

RESULT 9

US-09-850-716A-122
; Sequence 122, Application US/09850716A
; Patent No. US20020115139A1
; GENERAL INFORMATION:
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Retter, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C15
; CURRENT APPLICATION NUMBER: US/09/850,716A
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 122
; LENGTH: 1475
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-850-716A-122

Query Match 99.4%; Score 256.4; DB 10; Length 1475;
Best Local Similarity 99.6%; Pred. No. 8.6e-77;
Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAACCTGCTATGAGGGGAATGGTCACTTTTACCGAGAAAGGCCAGCACTGACACCATG 60
DB 282 AAAACCTGCTATGAGGGGAATGGTCACTTTTACCGAGAAAGGCCAGCACTGACACCATG 341
QY 61 GGCCGGCCCTGCGCTGCCCTGGAACCTCTGCCACTGTCTTCAGCAAAAGCTACCATGCCAC 120
DB 342 GGCCGGCCCTGCGCTGCCCTGGAACCTCTGCCACTGTCTTCAGCAAAAGCTACCATGCCAC 401
QY 121 AGATCTAATGCTCTTCAAGCTGGGCGGAAACATAATTAATCTGAGGAAGCCAGACAAC 180
DB 402 AGATCTGATGCTCTTCAAGCTGGGCGGAAACATAATTAATCTGAGGAAGCCAGACAAC 461
QY 181 CGGAGCGGACCCCTGGTGTCTATGTGAGAGTGGGCGCTAAAGCCGCTTGTCCAAGAGTGCATG 240
DB 462 CGGAGCGGACCCCTGGTGTCTATGTGAGAGTGGGCGCTAAAGCCGCTTGTCCAAGAGTGCATG 521
QY 241 GTGCATGACTGCGCAGAT 258
DB 522 GTGCATGACTGCGCAGAT 539

RESULT 10

US-09-897-778-122
; Sequence 122, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.

```

; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C16
; CURRENT APPLICATION NUMBER: US/09/897,778
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 122
; LENGTH: 1475
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-897-778-122

```

Query Match	99.48;	Score 256.4;	DB 10;	Length 1475;
Best Local Similarity	99.68;	Pred. No. 8.6e-77;		
Matches 257; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

QY	1	AAACCTGCTATGAGGGGAATGGTCACTTTTACCAGGAAAGCCAGCAGCTGCACCATG	60
Db	282	AAAACCTGCTATGAGGGGAATGGTCACTTTTACCAGGAAAGCCAGCAGCTGCACCATG	341
QY	61	GGCGGGCCCTGCGTCGCCCTGGAACCTGCCCAGCTGTCTCTACGCAAACGTACCATGCCAC	120
Db	342	GGCGGGCCCTGCGTCGCCCTGGAACCTGCCCAGCTGTCTCTACGCAAACGTACCATGCCAC	401
QY	121	AGATCTAATGCTCTTCAGCTGGGCGCTGGGGAAACATAATTACTGCAGGAACCCAGACAAC	180
Db	402	AGATCTGATGCTCTTCAGCTGGGCGCTGGGGAAACATAATTACTGCAGGAACCCAGACAAC	461
QY	181	CGGAGCGACCCCTGGTGCTATGTGCAAGTGGGCGCTTAAGCCGCTTGTCCAAGAGTGCATG	240
Db	462	CGGAGCGACCCCTGGTGCTATGTGCAAGTGGGCGCTTAAGCCGCTTGTCCAAGAGTGCATG	521
QY	241	GTCGATGACTGCCGACAT	258
Db	522	GTCGATGACTGCCGACAT	539

RESULT 11
US-09-735-705-123

```

: Sequence 123, Application US/09735705
: Patent No. US20020052329A1
: GENERAL INFORMATION:
: APPLICANT: Wang, Tongtong
: APPLICANT: Fan, Liqun
: APPLICANT: Kalos, Michael D.
: APPLICANT: Bangur, Chaitanya S.
: APPLICANT: Hosken, Nancy
: APPLICANT: Fanger, Gary R.
: APPLICANT: Li, Samuel X.
: APPLICANT: Wang, Aijun
: APPLICANT: Skeiky, Yasir A.W.
: APPLICANT: Henderson, Robert A.
: APPLICANT: McNeill, Patricia D.
: APPLICANT: Fanger, Neil
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
: TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
: FILE REFERENCE: 210121.455C14
: CURRENT APPLICATION NUMBER: US/09/735,705
: CURRENT FILING DATE: 2000-12-12
: NUMBER OF SEQ ID NOS: 419
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 123
: LENGTH: 2294
: TYPE: DNA
: ORGANISM: Homo sapien
US-09-735-705-123

```

Query Match	99.48;	Score 256.4;	DB 10;	Length 2294;
Best Local Similarity	99.68;	Pred. No. 1e-76;		
Matches 257; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

QY	1	AAAACTGCTATGAGGGGAATGGTCACTTTTACCAGGAAGGCCAGCACTGACACCATG	60
Db	268	AAAACTGCTATGAGGGGAATGGTCACTTTTACCAGGAAGGCCAGCACTGACACCATG	327
QY	61	GGCCGGCCCTGCGTCCCTGGAACCTGCGCACTGTCTTACAGAAACGTACCATGCCAC	120
Db	328	GGCCGGCCCTGCGTCCCTGGAACCTGCGCACTGTCTTACAGAAACGTACCATGCCAC	387
QY	121	AGATCTAATGCTCTTCAAGCTGGGCGCTGGGAAACATAATTACTGCAAGAACCCAGACAAC	180
Db	388	AGATCTGATGCTCTTCAAGCTGGGCGCTGGGAAACATAATTACTGCAAGAACCCAGACAAC	447
QY	181	CGGAGGCGACCTGCTGCTATGTGCAGTGGGCGCTAAAGCCGCTTGCCAAAGAGTGCAATG	240
Db	448	CGGAGGCGACCCCTGCTGCTATGTGCAGTGGGCGCTAAAGCCGCTTGCCAAAGAGTGCAATG	507
QY	241	GTGCATGACTGCCAGAT	258
Db	508	GTGCATGACTGCCAGAT	525

RESULT 12

```

US-09-850-716A-123
; Sequence 123, Application US/09850716A
; Patent No. US20020115139A1
; GENERAL INFORMATION:
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Retter, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS
; TITLE OF INVENTION: AND DIAGNOSIS OF LONG
; FILE REFERENCE: 210121.455C15
; CURRENT APPLICATION NUMBER: US/09/850,716A
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 123
; LENGTH: 2294
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-850-716A-123

```

Query M

	Query Match	99.48;	Score 256.4;	DB 10;	Length 2294;	
	Best Local Similarity	99.6%;	Pred. No. 1e-76;			
	Matches 257;	Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	1 AAAACCTGCTATGAGGGGAATGGTCACTTTTACCAGAGAAGGCCAGCAGTCAGCACCATG 60 					
Db	268 AAACCTGTCTATGAGGGGAATGGTCACTTTTACCAGAGAAGGCCAGCAGTCAGCACCATG 327 					
QY	61 GGCCGGCCCTGCCTGCCCTGGAACCTCGCCA CTGTCCCTTCAGCAAACGTACCATGCCAC 120 					
Db	328 GGGCGGCCCTGCCTGCCCTGGAACCTCGCCA CTGTCCCTTCAGCAAACGTACCATGCCAC 387 					
QY	121 AGATCTAATGCTCTT CAGCTGGGCGTGGGAAACATAATTACTGCAGAACCCAGACAAC 180 					
Db	388 AGATCTAATGCTCTT CAGCTGGGCGTGGGAAACATAATTACTGCAGAACCCAGACAAC 447 					
QY	181 CGGAGCGCACCCCTGTGCTATGTG CAGGTGGGCCCTTAAGCCGCTTGTC AAGA GTGCATG 240 					
Db	448 CGGAGCGCACCCCTGTGCTATGTG CAGGTGGGCCCTTAAGCCGCTTGTC AAGA GTGCATG 507 					
QY	241 GTGCATGACTGCCG CAGAT 258 					
Db	508 GTGCATGACTGCCG CAGAT 525 					

RESULT 13

; Sequence 123, Application US/09897778
 ; Patent No. US20020147143A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong

Tue Nov 12 18:06:52 2002

pct-us02-27855-2.rnpb

Page 6

```

: APPLICANT: Mainnerakis, Margarita
: APPLICANT: Fanger, Gary R.
: APPLICANT: Vedvick, Thomas S.
: APPLICANT: Carter, Darick
: APPLICANT: Watanabe, Yoshihiro
: APPLICANT: Henderson, Robert A.
: APPLICANT: Peckham, David W.
: APPLICANT: Fanger, Neil
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
: TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
: FILE REFERENCE: 210121.455C16
: CURRENT APPLICATION NUMBER: US/09/897,778
: CURRENT FILING DATE: 2001-06-28
: NUMBER OF SEQ ID NOS: 467
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 123
: LENGTH: 2294
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-897-778-123

```

	Query Match	99.48;	Score 256.4;	DB 10;	Length 2294;	
	Best Local Similarity	99.6%;	Pred. No. 1e-76;			
	Matches 257;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;	
QY	1 AAACCTGCTATGAGGGGAATGTCACCTTTACCAGAGAAGGCCACACTGCACCATG 60 					
Db	268 AAAACCTGCTATGAGGGGAATGTCACCTTTACCAGAGAAGGCCACACTGCACCATG 327 					
QY	61 GGCCGGCCCTGCGCTGCCCCGTGAACTCTGCCACTGTCTTTCAGCAAACGTACCATGCCAC 120 					
Db	328 GGCCTGGCCTGCGCTGCCCCGTGAACTCTGCCACTGTCTTTCAGCAAACGTACCATGCCAC 387 					
QY	121 AGATCTAATGCTCTTTCAGCTGGGCTGGGGAACAATAATTACTGCAGGAACCCAGACAAC 180 					
Db	388 AGATCTGATGCTCTTTCAGCTGGGCTGGGGAACAATAATTACTGCAGGAACCCAGACAAC 447 					
QY	181 CGGAGGCGACCCCTGTCCTATGTGCAGGTGGCCCTAAAGCCCTTGCCAAGAGTGCATG 240 					
Db	448 CCGAGGCGACCCCTGTCCTATGTGCAGGTGGCCCTAAAGCCCTTGCCAAGAGTGCATG 507 					
QY	241 GTGCATGACTGCCGAGAT 258 					
Db	508 GTGCATGACTGCCGAGAT 525 					

RESULT 14
US-09-954-456-552
: Sequence 552, Application US/09954456
: Patent No. US20020115057A1
: GENERAL INFORMATION:
: APPLICANT: Young, Paul
: TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Can
: TITLE OF INVENTION: Sets
: FILE REFERENCE: 689290-76
: CURRENT APPLICATION NUMBER: US/09/954,456
: CURRENT FILING DATE: 2001-09-18
: PRIOR APPLICATION NUMBER: US/60/233,617
: PRIOR FILING DATE: 2000-09-18
: PRIOR APPLICATION NUMBER: US/60/234,052
: PRIOR FILING DATE: 2000-09-20
: PRIOR APPLICATION NUMBER: US/60/234,923
: PRIOR FILING DATE: 2000-09-25
: PRIOR APPLICATION NUMBER: US/60/235,134
: PRIOR FILING DATE: 2000-09-25
: PRIOR APPLICATION NUMBER: US/60/235,637
: PRIOR FILING DATE: 2000-09-26
: PRIOR APPLICATION NUMBER: US/60/235,638
: PRIOR FILING DATE: 2000-09-26
: PRIOR APPLICATION NUMBER: US/60/235,711
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: US/60/235,720
: PRIOR FILING DATE: 2000-09-27

```

; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 552
; LENGTH: 2036
;
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-552

```

Query Match	Similarity	27.48;	Score 70.8;	DB 10;	length 2036;
Best Local	Similarity	55.28;	Pred. No. 3.4e-14;		
Matches 138;	Conservative	0;	Mismatches 112;	Indels 0;	Gaps 0;
QY	6	CTGCTATGAGGGGAATGGTCACTTTTACCGAGAAAGGCCACCACTGACACCATGGGCCG	65		
Db	858	CTGCTTCTTGGGGAACGGCACTGGGTACCGTGGCGTGGCCAGCACCTCAGCCTCGGGCCT	917		
QY	66	GCCCTGCCTGCCCTGGAACTCTGCCACTGTCTCTCAGCAACGTAACCATGCCCCAGATC	125		
Db	918	CAGCTGCTGCGCTGGAACTCCGATCTGCTCTACCAAGAGCTGCACGTGACTCCGTGGG	977		
QY	126	TAATGCTCTTCAGCTGGGCGCTGGGGAAACATATTAATTACTGCGAACCAGACAACCGAG	185		
Db	978	CGCCGCGGCCCTGCTGGGCTGGGCCCCCACCATGCTACTGCGGAATCCGACAATGACGA	1037		
QY	186	GCGACCCCTGGTCTATGTGCAGAGTGGGCGCTAAAGCCCGCTTGTCCAGAAGTGCATGGTCA	245		
Db	1038	GAGGCCCTGGTCTACGTGTTGAAGACACGCGCGCTTCCTGGGAGTACTGCCGCGCTGGA	1097		
QY	246	TGACTGCGCA	255		
Db	1098	GGCCTGCGAA	1107		

```

RESULT 15
US-09-880-107-1612
; Sequence 1612, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1612
; LENGTH: 2036
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 D14012
US-09-880-107-1612

Query Match      27.4%; Score 70.8; DB 10; Length 2036;
Best Local Similarity 55.2%; Pred. No. 3.4e-14;
Matches 138; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

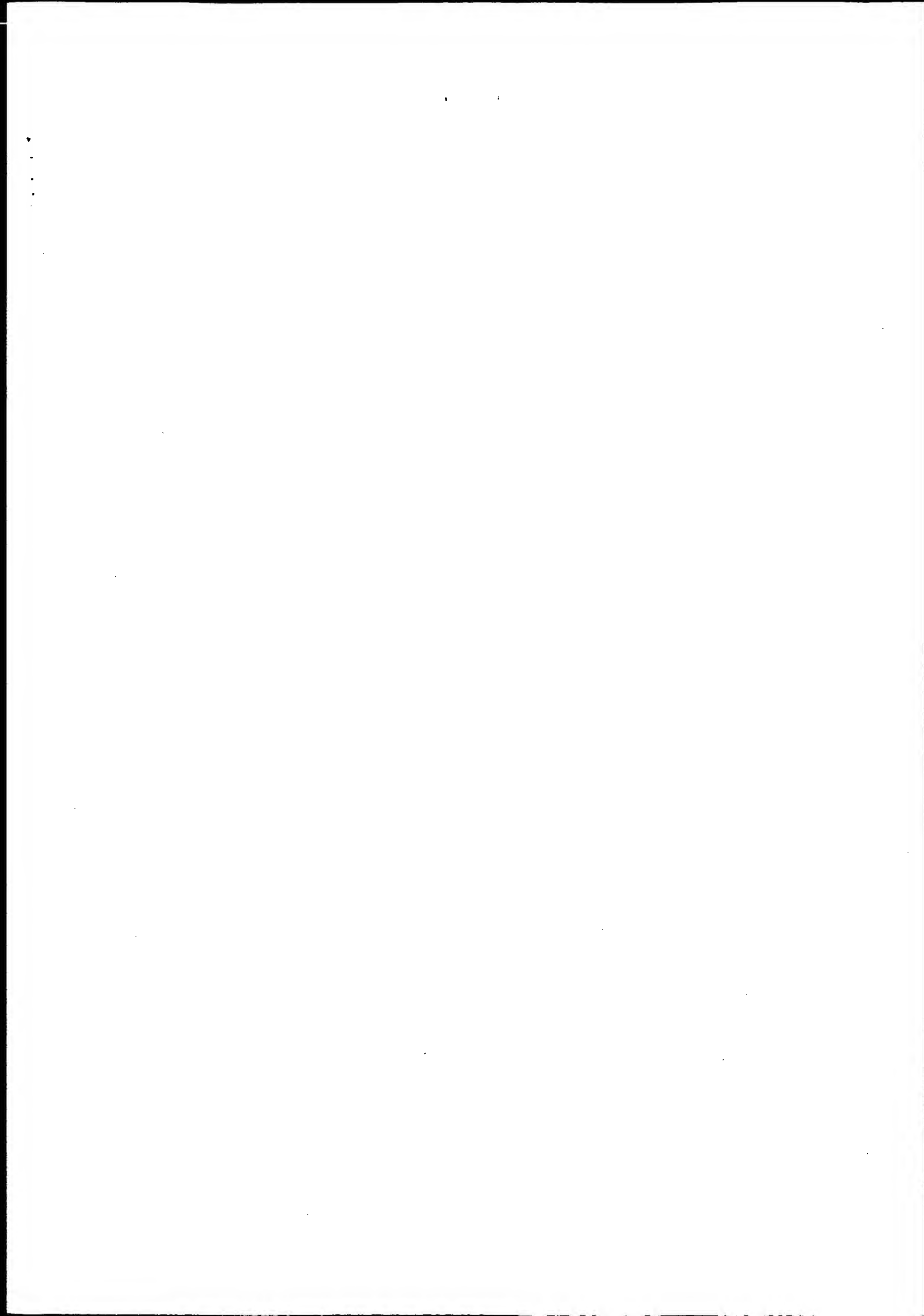
QY      6 CTGCTATGAGGGGAATGTCACCTTTTACCGAGGAAGGCCAGACACTGACACCATGGGCGG 65
      ||||| ||||| || ||||| || ||||| || ||||| || |||||
Db      858 CTGCTTCTTTGGGAACGGCACCTGGGTACCGTGGCGTGCCAGCACCTCAGCCCTCGGGCCT 917
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Y      66 GCCCTGCCTGCCTCGAACCTCTGCCACACTGTCTCTTCAGCAAAACGTATCCATGCCACAGATC 125
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```
Db 918 CAGCTGCTGGCTGGAAGTCCGATCTGCTTACAGAGAGCTGCAGCTGGACTCCGTGGG 977
QY 126 TAATGCTCTTCAGCTGGGCTGGGAAACATATATTACTGCAGAGAACCCAGACACCGGAG 185
Db 978 CGCCGGGGCTGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 1037
QY 186 GCGACCTGCTGCTATGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGG 245
Db 1038 GAGGCCCTGCTGCTACGTCGTAAGAGACGCGCTCTCTGGAAGTACTGCCGCTGGA 1097
QY 246 TGAATGCGCA 255
Db 1098 GGCTGCGAA 1107
```

Search completed: November 11, 2002, 12:58:35
Job time : 46 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 11, 2002, 11:36:26 ; Search time 1968 Seconds
(without alignments)
2123.189 Million cell updates/sec

Title: PCT-US02-27855-2

Perfect score: 258
Sequence: 1 aaacactgcatgaggggaa.....tggtgcatgactgcagat 258

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 809774376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	256.4	99.4	700	12	BE877476 601485319
2	256.4	99.4	767	12	BE877571 601486523
3	256.4	99.4	799	9	AU131687 AU131687
4	256.4	99.4	870	14	BQ641748 AGENCOURT
5	256.4	99.4	871	12	BG748424 602705991
6	256.4	99.4	913	14	BQ886981 AGENCOURT

7	256.4	99.4	930	12	BE742186	BE742186 601575784
8	256.4	99.4	932	14	BQ961887	BQ961887 AGENCOURT
9	256.4	99.4	963	14	BQ929215	BQ929215 AGENCOURT
10	256.4	99.4	1001	14	BQ922207	BQ922207 AGENCOURT
11	256.4	99.4	1063	14	BM903637	BM903637 AGENCOURT
12	256.4	99.4	1067	14	BM920114	BM920114 AGENCOURT
13	256.4	99.4	1075	12	BE742275	BE742275 601575693
14	256.4	99.4	1087	12	BE897378	BE897378 601437369
15	256.4	99.4	1164	12	BF568088	BF568088 602184082
16	255.4	99.0	715	9	AU135252	AU135252 AU135252
17	254.8	98.8	479	14	BM843656	BM843656 K-EST0121
18	254.8	98.8	602	14	BM849144	BM849144 K-EST0129
19	254.8	98.8	647	12	BE889213	BE889213 601513488
20	254.8	98.8	980	14	BQ920793	BQ920793 AGENCOURT
21	253.8	98.4	595	14	BM853271	BM853271 K-EST0134
22	248.4	96.3	474	12	BF873834	BF873834 IL3-ET011
23	245.4	95.1	718	9	AU128113	AU128113 AU128113
24	245.4	95.1	822	13	BI833559	BI833559 603088241
25	244.4	94.7	1130	12	BG832239	BG832239 602765624
26	243.8	94.5	913	14	BQ887418	BQ887418 AGENCOURT
27	240.6	93.3	775	13	BM048314	BM048314 603625527
28	238	92.2	710	12	BG331244	BG331244 602431960
29	234.4	90.9	625	12	BE880196	BE880196 601491257
30	233.4	90.5	928	12	BG330224	BG330224 602429932
31	232.8	90.2	803	12	BG741013	BG741013 602634740
32	230.4	89.3	672	12	BG254622	BG254622 602368575
33	226.8	87.9	549	14	BM848159	BM848159 K-EST0127
34	224.8	87.1	261	14	BM761640	BM761640 K-EST0042
35	221.6	85.9	1093	12	BE874867	BE874867 601488917
36	220.6	85.5	1152	13	BI759308	BI759308 603042953
37	203.6	78.9	952	12	BG035185	BG035185 602324887
38	198.4	76.9	223	9	AA158183	AA158183 z055d07.r
39	191.6	74.3	638	10	BB666464	BB666464 BB666464
40	191.6	74.3	655	10	BB555211	BB555211 BB555211
41	165.4	64.1	337	12	BG007464	BG007464 RC2-GN025
42	148.6	57.6	472	14	BM847321	BM847321 K-EST0126
43	142.4	55.2	901	14	BQ420628	BQ420628 AGENCOURT
44	135.4	52.5	420	10	AW238819	AW238819 xb34f12.y
45	134.4	52.1	488	9	AL549674	AL549674 AL549674

ALIGNMENTS

RESULT 1
BE877476 700 bp mRNA linear EST 20-OCT-2000
LOCUS 601485319F1 NIH_MGC_69 Homo sapiens cDNA clone IMAGE:3887620 5',
DEFINITION mRNA sequence.
ACCESSION BE877476
VERSION BE877476.1 GI:10326252
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 700)
AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP/Gazdar
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLM9666 row: f column: 05
High quality sequence stop: 671.
Location/Qualifiers 1..700

FEATURES
source

Query Match	99.48;	Score 256.4;	DB 9;	Length 799;
Best Local Similarity	99.68;	Pred. No. 5.2e-67;		
Matches 257; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

[illegible]

RESULT 4	
BQ641748	
LOCUS	870 bp
DEFINITION	mRNA
	linear
	EST 15-JUL-2002
BQ641748	
AGENCOURT_8286901	NIH_MGC_43
	Homo sapiens
	CDNA clone IMAGE:6292200
5', mRNA sequence.	

ACCESSION	BQ641748	
VERSION	BQ641748.1	GI:21765920
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
JOURNAL 1 (bases 1 to 870)
COMMENT NIH-MGC <http://mgc.ncl.nih.gov/>.
CONTACT National Institutes of Health, Mammalian Gene Collection (MGC)
UNPUBLISHED (1999)
CONTACT: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://limage.llnl.gov>
Plate: LNCM2493 row: e column: 01
High quality sequence stop: 677.

FEATURES

Location/Qualifiers
1. .870

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6292200"
/clone_11b="NIH_MGC_43"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: eye; Vector: pOTB7; Site_1: xhoI; Site_2:
EcoRI: cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/xhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library. |"

```

BASE COUNT
ORIGIN

Query Match

99.48; Score 256.4; DB 14; Length 870;

Best Local Similarity 99.68; Pred. NO. 5.4e-67;
Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	AAACCTGCTATGAGGGGAATGGTACTTTTACCGAGGAAGGCCAGCACTGACACCATG	60
DB	128	AAACCTGCTATGAGGGGAATGGTACTTTTACCGAGGAAGGCCAGCACTGACACCATG	187
QY	61	GGCCGGCCCTGCGCTGCCCTGGAACCTGCCCAGCTGTCCTTACAGCAAACTGACCATGCCAC	120
DB	188	GGCCGGCCCTGCGCTGCCCTGGAACCTGCCCAGCTGTCCTTACAGCAAACTGACCATGCCAC	247
QY	121	AGATCTAATGCTCTTCACTGGCGCTGGGAAACATAATTAATCTGAGGAAGCCAGACAAAC	180
DB	248	AGATCTGATGCTCTTCACTGGCGCTGGGAAACATAATTAATCTGAGGAAGCCAGACAAAC	307
QY	181	CGGAGGCGACCCCTGGTGTCTATGTGCAGGTGGGCTAAAGCCGCTTGTCCAAGAGTGCATG	240
DB	308	CGGAGGCGACCCCTGGTGTCTATGTGCAGGTGGGCTAAAGCCGCTTGTCCAAGAGTGCATG	367
QY	241	GTGCATGACTGCCGACAT	258
DB	368	GTGCATGACTGCCGACAT	385

RESULT 5					
BG748424					
LOCUS	871 bp	mRNA	linear	EST 15-MAY-2001	
DEFINITION	602705991F1 NIH_MGC_43	Homo sapiens	CDNA clone	IMAGE:4842623	5',
	mRNA sequence.				

ACCESSION	BG748424	GI:14059077
VERSION	BG748424.1	
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 871)
NIH-MGC <http://mgc.ncl.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubln Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L16M1677 row: e column: 24
High quality sequence stop: 802.

FEATURES

Location/Qualifiers
1. .871

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4842623"
/clone_11b="NIH_MGC_43"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
Note: this is a NIH_MGC Library. |"

```

BASE COUNT
ORIGIN

Query Match

```

Best Local Similarity: 99.00; Read: 0.40 0.77
Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```


QY	1	AAACCTGCTATGAGGGGAATGGTCACTTTTACCAGAGAAAAGGCCAGCACTGACACCATG	60
Db	308	AAAACTGCTATGAGGGGAATGGTCACTTTTACCAGAGAAAAGGCCAGCACTGACACCATG	367
QY	61	GCGCCGCCCTGCCTGCCCCGTGGAACCTGCCCACCTGTCTTCAGACAACGTACCATGCCAC	120
Db	368	GCGCCGCCCTGCCTGCCCCGTGGAACCTGCCCACCTGTCTTCAGACAACGTACCATGCCAC	427
QY	121	AGATCTAATGCTCTTCAAGCTGGGCTTGGGGAAACATAATTACTGCAGSAACCCAGACAAC	180
Db	428	AGATCTGATGCTCTTCAAGCTGGGCTTGGGGAAACATAATTACTGCAGSAACCCAGACAAC	487
QY	181	CGGAGGCGAACCCCTGGTGCTATATGTCAGGTGGGCCCTAAAGCCGCTTGTCCAAGAGTGCATG	240
Db	488	CGGAGGCGAACCCCTGGTGCTATATGTCAGGTGGGCCCTAAAGCCGCTTGTCCAAGAGTGCATG	547
QY	241	GTGCATGACTGCGCAGAT	258
Db	548	GTGCATGACTGCGCAGAT	565

RESULT	6
BQ886981	
LOCUS	
DEFINITION	BQ886981 913 bp mRNA linear EST 16-AUG-2002
ACCESSION	AGENCOURT_8675606 NIH_MGC_40 Homo sapiens cdna clone IMAGE:6380053
VERSION	5' , mRNA sequence .
KEYWORDS	BQ886981
SOURCE	BQ886981.1 GI:22278995
	EST.
	human .

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 913)
NIH-MGC <http://mgc.ncl.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM2568 row: a column: 14
High quality sequence stop: 595.
location/Qualifiers
1. .913

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6380053"
/clone_lib="NIH_MGC_40"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

```

Query Match	99.48;	Score 256.4;	DB 14;	Length 913;
Best Local Similarity	99.6%;	Pred. NO. 5.5e-67;		
Matches 257;	Conservative	0;	Mismatches 1;	Indels 0;
Matches				Gaps 0;
1	AAAACCTGCTATGAGGGCAATGCTCACTTTTACCGAGGAAGGCCAGCACTGCACCATG	60		

Db	310	AAACCTGCTATGAGGGGAATGGTCACTTTTACCAGGAAAGCCAGCACTGACACCACTG	369
QY	61	GGCCGGCCCTGCCCTGCCCTTGGAACTCTGCCACTGTCTCTTACAGCAAAACGTACCATGCCAC	120
Db	370	GGCCGGCCCTGCCCTGCCCTTGGAACTCTGCCACTGTCTCTTACAGCAAAACGTACCATGCCAC	429
QY	121	AGATCTAATGCTCTTCAAGCTGGGCGCTGGGGAAACATAATTACTGCAGGAACCCAGAACAC	180
Db	430	AGATCTGATGCTCTTCAAGCTGGGCGCTGGGGAAACATAATTACTGCAGGAACCCAGAACAC	489
QY	181	CGGAGGCGACCCCTGCTGCTATGTGACAGTGGGCGCTAAAGCCGCTGTGCCAAGATGCATG	240
Db	490	CGGAGGCGACCCCTGCTGCTATGTGACAGTGGGCGCTAAAGCCGCTGTGCCAAGATGCATG	549
QY	241	GTGCATGACTGCCGAGAT	258
Db	550	GTGCATGACTGCCGAGAT	567

RESULT	7
LOCUS	BE742186
DEFINITION	930 bp mRNA linear EST 15-SEP-2000 601575784F1 NIH_MGC_9 Homo sapiens CDNA clone IMAGE:3836612 5'
ACCESSION	BE742186
VERSION	BE742186.1 GI:10156178
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 930)
NIH-MGC <http://mgc.ncl.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC),
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

Email: cqapbs-r@mail.nih.gov
 Tissue Procurement: DCTD/DTF
 CDNA Library Preparation: Ling Hong/Rubln Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: L1CM520 row: h column: 21
 High quality sequence stop: 743.
 Location/Qualifiers
 1. .930

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3836612"
/clone_lib="NIH_MGC_9"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: ovary; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAC(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

```

	Query Match	99.4%;	Score 256.4;	DB 12;	Length 930;
	Best Local Similarity	99.6%;	Pred. No. 5,6e-67;		
	Matches 257;	Conservative	0;	Mismatches 1;	Indels 0;
QY	1 AAACCTGCTATGAGGGGAATGTGTCACCTTTACCGAGGAAGGCCAGCACTGCACCATG	60			
D0	20 AAACCTGCTATGAGGGGAATGTGTCACCTTTACCGAGGAAGGCCAGCACTGCACCATG	79			

QY 61 GCGCGGCCCTGCTGCCCTGGAACTCTGCCACTGTCCTTCAGCAAAAGCTACATGCCAC 120
|||||
Db 80 GCGCGGCCCTGCTGCCCTGGAACTCTGCCACTGTCCTTCAGCAAAAGCTACATGCCAC 139
QY 121 AGATCTAATGCTCTTCAGCTGGGCGCTGGGAAACATTAATTAATGAGAACCCAGACAC 180
|||||
Db 140 AGATCTGATGCTCTTCAGCTGGGCGCTGGGAAACATTAATTAATGAGAACCCAGACAC 199
QY 181 CGGAGCGACCCCTGCTGCTATGTGACAGTGGGCGCTAAAGCCGCTGTCCAGAGTGCATG 240
|||||
Db 200 CGGAGCGACCCCTGCTGCTATGTGACAGTGGGCGCTAAAGCCGCTGTCCAGAGTGCATG 259
QY 241 GTGCATGACTGGCGAGAT 258
|||||
Db 260 GTGCATGACTGGCGAGAT 277

RESULT 8
BQ961887 932 bp mRNA linear EST 21-AUG-2002
LOCUS
DEFINITION AGENCOURT_8778892 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:6378687
5', mRNA sequence.
ACCESSION BQ961887
VERSION BQ961887.1 GI:223777365
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 932)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2564 row: h column: 16
High quality sequence stop: 606.
location/Qualifiers
1. 932
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6378687"
/clone_lib="NIH_MGC_42"
/tissue_type="epithelioid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library. 1"
BASE COUNT 225 a 281 c 248 g 177 t 1 others
ORIGIN

Query Match 99.4%; Score 256.4; DB 14; Length 932;
Best Local Similarity 99.6%; Pred. No. 5.6e-67;
Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAACCTGCTATGAGGGGAATGTCACCTTTACCGAGGAAAGGCCAGCACTGACACCATG 60
|||||
Db 303 AAACCTGCTATGAGGGGAATGTCACCTTTACCGAGGAAAGGCCAGCACTGACACCATG 362
QY 61 GCGCGGCCCTGCTGCCCTGGAACTCTGCCACTGTCCTTCAGCAAAAGCTACATGCCAC 120

Db 363 GCGCGGCCCTGCTGCCCTGGAACTCTGCCACTGTCCTTCAGCAAAAGCTACATGCCAC 422
QY 121 AGATCTAATGCTCTTCAGCTGGGCGCTGGGAAACATTAATTAATGAGAACCCAGACAC 180
|||||
Db 423 AGATCTGATGCTCTTCAGCTGGGCGCTGGGAAACATTAATTAATGAGAACCCAGACAC 482
QY 181 CGGAGCGACCCCTGCTGCTATGTGACAGTGGGCGCTAAAGCCGCTGTCCAGAGTGCATG 240
|||||
Db 483 CGGAGCGACCCCTGCTGCTATGTGACAGTGGGCGCTAAAGCCGCTGTCCAGAGTGCATG 542
QY 241 GTGCATGACTGGCGAGAT 258
|||||
Db 543 GTGCATGACTGGCGAGAT 560

RESULT 9
BQ929215 963 bp mRNA linear EST 20-AUG-2002
LOCUS
DEFINITION AGENCOURT_8958654 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6460542
5', mRNA sequence.
ACCESSION BQ929215
VERSION BQ929215.1 GI:22344246
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 963)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2647 row: g column: 07
High quality sequence stop: 566.
location/Qualifiers
1. 963
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6460542"
/clone_lib="NIH_MGC_101"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site_1: EcoRI; Site_2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library. 1"
BASE COUNT 221 a 301 c 256 g 182 t 3 others
ORIGIN

Query Match 99.4%; Score 256.4; DB 14; Length 963;
Best Local Similarity 99.6%; Pred. No. 5.7e-67;
Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAACCTGCTATGAGGGGAATGTCACCTTTACCGAGGAAAGGCCAGCACTGACACCATG 60
|||||
Db 277 AAACCTGCTATGAGGGGAATGTCACCTTTACCGAGGAAAGGCCAGCACTGACACCATG 336
QY 61 GCGCGGCCCTGCTGCCCTGGAACTCTGCCACTGTCCTTCAGCAAAAGCTACATGCCAC 120
|||||
Db 337 GCGCGGCCCTGCTGCCCTGGAACTCTGCCACTGTCCTTCAGCAAAAGCTACATGCCAC 396

QY 121 AGATCTAATGCTCTTCAGCTGGGGCCGTGGGGAACATTAATTACTGCAGGAACCCAGACAC 180
||||| |||||||
Db 397 AGATCTGATGCTCTTCAGCTGGGGCCGTGGGGAACATTAATTACTGCAGGAACCCAGACAC 456
QY 181 CGGAGCGGACCCCTGGTGTCTATGTGCAGGTGGGCTTAAGCCGCTTGTCCAGAGTGCATG 240
||||| |||||||
Db 457 CGGAGCGGACCCCTGGTGTCTATGTGCAGGTGGGCTTAAGCCGCTTGTCCAGAGTGCATG 516
QY 241 GTGCATGACTGCGCAGAT 258
||||| |||||||
Db 517 GTGCATGACTGCGCAGAT 534

RESULT 10
BO922207 1001 bp mRNA linear EST 20-AUG-2002
LOCUS
DEFINITION AGENCOURT_8945721 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6459780
5', mRNA sequence.
ACCESSION BO922207
VERSION BO922207.1 GI:22337238
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1001)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM2645 row: g column: 13
High quality sequence stop: 520.

FEATURES

source
1.1001
Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6459780"
/clone_lib="NIH_MGC_101"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site_1: EcoRI; Site_2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

BASE COUNT 230 a 306 c 272 g 192 t 1 others
ORIGIN

Query Match 99.4%; Score 256.4; DB 14; Length 1001;
Best Local Similarity 99.6%; Pred. No. 5.8e-67;
Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAACCTGCTATGAGGGGAATGTCACCTTTTACCGAGGAAGGCCAGCACTGACACCATG 60
||||| |||||||
Db 277 AAAACCTGCTATGAGGGGAATGTCACCTTTTACCGAGGAAGGCCAGCACTGACACCATG 336
QY 61 GGGCGGGCCCTGCGCCCTGGAACCTCTGCCACTGTCTTCAGCAAAACGTACCATGCCAC 120
||||| |||||||
Db 337 GGGCGGGCCCTGCGCCCTGGAACCTCTGCCACTGTCTTCAGCAAAACGTACCATGCCAC 396
QY 121 AGATCTAATGCTCTTCAGCTGGGGCCGTGGGGAACATTAATTACTGCAGGAACCCAGACAC 180

Db 397 AGATCTGATGCTCTTCAGCTGGGGCCGTGGGGAACATTAATTACTGCAGGAACCCAGACAC 456
||||| |||||||
QY 181 CGGAGCGGACCCCTGGTGTCTATGTGCAGGTGGGCTTAAGCCGCTTGTCCAGAGTGCATG 240
||||| |||||||
Db 457 CGGAGCGGACCCCTGGTGTCTATGTGCAGGTGGGCTTAAGCCGCTTGTCCAGAGTGCATG 516
QY 241 GTGCATGACTGCGCAGAT 258
||||| |||||||
Db 517 GTGCATGACTGCGCAGAT 534

RESULT 11
BM903637 1063 bp mRNA linear EST 12-MAR-2002
LOCUS
DEFINITION AGENCOURT_6622158 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5722049
5', mRNA sequence.
ACCESSION BM903637
VERSION BM903637.1 GI:19353039
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1063)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Invitrogen
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12707 row: h column: 18
High quality sequence stop: 742.

FEATURES

source
1.1063
Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5722049"
/clone_lib="NIH_MGC_125"
/lab_host="DH10B"
/note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;
Site_1: EcoRV (destroyed); Site_2: NotI; RNA source pool
of three ovaries, from females ranging in age from 38 to
49 yo. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 2.1 kb, insert size range 1-3.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 036."

BASE COUNT 257 a 315 c 282 g 208 t 1 others
ORIGIN

Query Match 99.4%; Score 256.4; DB 14; Length 1063;
Best Local Similarity 99.6%; Pred. No. 6e-67;
Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAACCTGCTATGAGGGGAATGTCACCTTTTACCGAGGAAGGCCAGCACTGACACCATG 60
||||| |||||||
Db 311 AAAACCTGCTATGAGGGGAATGTCACCTTTTACCGAGGAAGGCCAGCACTGACACCATG 370
QY 61 GGGCGGGCCCTGCGCCCTGGAACCTCTGCCACTGTCTTCAGCAAAACGTACCATGCCAC 120
||||| |||||||
Db 371 GGGCGGGCCCTGCGCCCTGGAACCTCTGCCACTGTCTTCAGCAAAACGTACCATGCCAC 430
QY 121 AGATCTAATGCTCTTCAGCTGGGGCCGTGGGGAACATTAATTACTGCAGGAACCCAGACAC 180
||||| |||||||
Db 431 AGATCTGATGCTCTTCAGCTGGGGCCGTGGGGAACATTAATTACTGCAGGAACCCAGACAC 490

QY 181 CGAGGCGACCTGTGCTATGTGCAGGTGGCCCTAAAGCCGCTTGCCAGAGTGCATG 240
|||||
Db 491 CGAGGCGACCTGTGCTATGTGCAGGTGGCCCTAAAGCCGCTTGCCAGAGTGCATG 550
QY 241 GTGCATGACTGCGCAGAT 258
|||||
Db 551 GTGCATGACTGCGCAGAT 568

RESULT 12

BM920114 1067 bp mRNA linear EST 12-MAR-2002
LOCUS AGENCOURT_6706522 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5749987
DEFINITION 5', mRNA sequence.
ACCESSION BM920114
VERSION BM920114.1 GI:19370493
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1067)
NIH-MGC http://mgi.ncl.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM12780 row: d column: 20
High quality sequence stop: 754.

FEATURES

Source

1.1067
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5749987"
/clone_lib="NIH_MGC_120"
/lab_host="DH10B"
/note="Organ: pooled pancreas and spleen; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of spleen and pancreas from 28 yo
male. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 1.5 kb, insert size range 1-2.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 025. Note: this is a NIH_MGC Library."

BASE COUNT 256 a 323 c 273 g 208 t 7 others

ORIGIN

Query Match 99.4%; Score 256.4; DB 14; Length 1067;
Best Local Similarity 99.6%; Pred. No. 6e-67;
Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAACCTGCTATGAGGGGAATGCTACTTTTACCGAGAAAGCCAGCACTGACACCATG 60
|||||
Db 328 AAAACCTGCTATGAGGGGAATGCTACTTTTACCGAGAAAGCCAGCACTGACACCATG 387
61 GGCCGGCCCTGCTGCTGAGCACTGCTCTTACGCAAAAGTACCAGCCAC 120
|||||
Db 388 GGCCGGCCCTGCTGCTGAGCACTGCTCTTACGCAAAAGTACCAGCCAC 447
121 AGATCTATGCTCTTACGCTGGCCCTGGGGAACATATTAATGACAGAACCCAGACAAC 180
|||||
Db 448 AGATCTGATGCTCTTACGCTGGCCCTGGGGAACATATTAATGACAGAACCCAGACAAC 507
181 CGGAGGCGACCTGTGCTATGTGCAGGTGGCCCTAAAGCCGCTTGCCAGAGTGCATG 240

Db 508 CGAGGCGACCTGTGCTATGTGCAGGTGGCCCTAAAGCCGCTTGCCAGAGTGCATG 567
|||||
QY 241 GTGCATGACTGCGCAGAT 258
|||||
Db 568 GTGCATGACTGCGCAGAT 585

RESULT 13

BE742275 1075 bp mRNA linear EST 15-SEP-2000
LOCUS 601575693F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3836663 5',
DEFINITION mRNA sequence.
ACCESSION BE742275
VERSION BE742275.1 GI:10156267
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1075)
NIH-MGC http://mgi.ncl.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at: image.llnl.gov
Plate: LLM520 row: j column: 24
High quality sequence stop: 681.

FEATURES

Source

1.1075
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3836663"
/clone_lib="NIH_MGC_9"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: ovary; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 226 a 322 c 290 g 237 t

ORIGIN

Query Match 99.4%; Score 256.4; DB 12; Length 1075;
Best Local Similarity 99.6%; Pred. No. 6e-67;
Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAACCTGCTATGAGGGGAATGCTACTTTTACCGAGAAAGCCAGCACTGACACCATG 60
|||||
Db 20 AAAACCTGCTATGAGGGGAATGCTACTTTTACCGAGAAAGCCAGCACTGACACCATG 79
61 GGCCGGCCCTGCTGCTGAGCACTGCTCTTACGCAAAAGTACCAGCCAC 120
|||||
Db 80 GGCCGGCCCTGCTGCTGAGCACTGCTCTTACGCAAAAGTACCAGCCAC 139
121 AGATCTATGCTCTTACGCTGGCCCTGGGGAACATTAATTAATGACAGAACCCAGACAAC 180
|||||
Db 140 AGATCTGATGCTCTTACGCTGGCCCTGGGGAACATTAATTAATGACAGAACCCAGACAAC 199
181 CGGAGGCGACCTGTGCTATGTGCAGGTGGCCCTAAAGCCGCTTGCCAGAGTGCATG 240
|||||
Db 200 CGGAGGCGACCTGTGCTATGTGCAGGTGGCCCTAAAGCCGCTTGCCAGAGTGCATG 259

